INDUCTIVE CONCEPT FORMATION IN NORMAL AND RETARDED SUBJECTS

Cooperative Research Project No. 833(8520)

CLAUDE B. ELAM

Texas Christian University

q153.1 E37i 63-02978
Elam
Inductive concept formation in
normal and retarded subjects

INDUCTIVE CONCEPT FORMATION IN NORMAL AND RETARDED SUBJECTS

Cooperative Research Project No. 833(8520)

Claude B. Elam

Texas Christian University
Fort Worth, Texas

1962

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

ABSTRACT

Concept formation, as it is affected by subjects (normal and retarded), problems (simultaneous, successive, and similarity-difference), responses (manual and verbal), reinforcement (100% and partial reinforcement), and stimuli was investigated. Eleven separate experiments are described.

It was found (1) concept attainment is facilitated by the verbal response, (2) two concepts can develop simultaneously, (3) concept of race is more easily utilized than age or sex, (4) amount of reinforcement procedure favors normal subjects while probability of reinforcement procedure favors retarded subjects, and (5) stimulus synthesis can be described as proportional to the strength of the various cues given.

TABLE OF CONTENTS

Section		Page
1	Introduction	6
2	Effect of Type of Response on the Solution of the Similarity-Difference Problem	10
3	Comparison of Four Grade Levels on the Similarity-Difference Problem	18
4	Concept Reversal Under Two Problem Conditions	31
5	Comparison of Normal and Retarded Children on the Similarity-Difference Problem	43
6	Comparison of Two Reinforcement Schedules on Problem Reversal	57
7	The Effect of Varying the Probability and the Amount of Reinforcement in Normal and Retarded Subjects	61
8	Effect of Partial Reinforcement on Normal and Retarded Subjects	66
9	Concept Formation in the Solution of a Three Light Problem	69
10	Comparison of Three Methods of Reinforcement Quantification	74
11	Perceptual Synthesis Involving Two Stimulus Variations	87
12	Stimulus Synthesis Involving Four Stimulus Variations	94
13	General Summary	105
	References	110

LIST OF TABLES

Table	Section	Title	Page
1	2	Stimulus-Response Relationship in the Similarity-Difference Problem	10
2	2	Mean Correct Responses	13
3	2	Analysis of Variance	14
4	2	Statistical Comparison of Response Treatments	15
5	2	Mean Response Time	15
6	3	Description of Stimuli	20
7	3	Mean Trials to Criterion	21
8	3	Analysis of Trials to Criterion	21
9	3	$\underline{t}$ Scores on Trials to Criterion	22
10	3	Mean Errors: Phase 1	22
11	3	Analysis of Errors: Phase 1	23
12	3	$\underline{t}$ Scores: Phase 1	23
13	3	Subject Assignment: Phase 2	25
14	3	Means for Trials to Criterion: Phase 2	26
15	3	Analysis of Variance on Trials: Phase 2	27
16	3	Mean Errors: Phase 2	28
17	3	Analysis of Variance on Errors: Phase 2	29
18	4	Stimulus-Response Relationships in the Simultaneous, Successive, and Similarity-Difference Problem	31
19	4	Correct Stimulus-Response Relationships	35
20	4	Experimental Design	36-37
21	4	Mean Errors Based on Groups	39
22	4	Mean Errors Based on Treatments	40
23	4	Analysis of Variance: Stage 3	41
24	5	Stimulus Variations Used	45

Table	Section	Title	Page
25	5	Mean Correct Responses: Stage 1	47
26	5	Analysis of Subject Source and Number of Stimulus Variations: Stage 1	48
27	5	Analysis of 3V Data: Stage 1	49
28	5	Comparison by Groups of 3V Data: Stage 1	50-51
29	5	Analysis of 4V Data: Stage 1	52
30	5	Comparison by Treatments of 4V Data: Stage 1	52
31	5	Mean Correct Responses: Stage 3	53
32	5	Analysis of Subject Source and Stimulus Variation: Stage 3	53
33	5	Analysis of 3V Data: Stage 3	54
34	5	Analysis of 4V Data: Stage 3	55
35	5	Comparison of Grade Level: Stage 3	55
36	6	Example of Problem	58
37	7	Subject Treatment Groups	62
38	7	Reinforcement Schedule	63
39	7	Mean Correct Responses	64
40	7	Analysis of Reinforcement and Subject Source Variance	64
41	7	Results of _t_ Tests	65
42	8	Means and Group Sizes	67
43	8	Analysis of Reinforcement and Subject Source Variance	68
44	9	Description of Problems	70
45	9	Comparison of Problem Differences	71
46	9	Analysis of Problem, Subject, and Order of Presentation Variances	73
47	10	Subject Treatment Groupings	76

Table	Section	Title	Page
48	10	Mean Correct Responses for each Block of 42 Trials	84
49	10	Summary of the Analyses of Variance	85
50	11	Correlations between Predicted and Actual Outcome	90
51	11	Significance of the Differences between Correlations	92
52	12	Analysis of Variance Based on Correct Responses	99
53	12	Mean Correct Responses	100
54	12	$\underline{t}$ Scores Based on Group Comparisons	100
55	12	Correlations between Predicted and Actual Outcome	101
56	12	Significance of Differences between Correlations	103

LIST OF FIGURES

Figure	Section	Title	Page
1	4	Example of Stimuli	34
2	10	Diagramatic Representation of the Problem	75
3	10	Frequency of X as a Function of the Variation in A	78
4	10	Frequency of Y as a Function of the Variation in A	78
5	10	Frequency of X as a Function of the Variation in B	79
6	11	P_A as a Function of the Variation in A	90
7	12	Frequency of X as a Function of the Variation in A	96
8	12	Frequency of X as a Function of the Variation in B	96
9	12	Frequency of X as a Function of the Variation in C	97
10	12	Frequency of X as a Function of the Variation in D	97

SECTION 1

INTRODUCTION

The term "concept formation" has been used in numerous connections. For some (5) its meaning is indistinguishable from discriminative learning. Others (7,17,18) prefer a more restricting definition. It is not the purpose of the present report to select between definitions. The tendency here, however, is to favor a rather broad use of the term.

By not recognizing the ubiquity of the phenomena, a definition of concept formation runs the risk of being descriptive of the method of measurement rather than of the process measured. It could be said, for example, that the study of concept formation deals only with abstract relations such as number, relative position, size, etc., and that object classes such as books, men, and chairs are more properly the stimulus material of discriminative learning. Notwithstanding the familiar distinction between the abstract and the concrete, there is insufficient reason to exclude such classes as cow or book from consideration. An object is a synthesis of its attributes. A cow is simply a very cow-like object, and some cows are more cow-like than others. In any case, the convention of distinguishing between things and the relationship between things is not always recognized by children, and the process by which cows are assembled into a class does not differ from the process by which "many" is distinguished from "few."

Concepts relate to the manner in which a person organizes his stimulus universe into perceptual and response categories. Even to perceive a stimulus as an object is to classify it. An approach based upon this definition cuts across the more familiar areas of learning and perception.

Few escape the error of assuming that classes are somehow predesignated and immutable. The "nature" of an object or the relationship between objects depends, however, on the use they are put to. In one situation a book, to a literate adult, will fall into a conceptual class which also includes magazines and newspapers; in a different situation it will be conjoined with a host of objects, any of which will serve to raise a child's seat higher to the table. For the child it will more appropriately be related to building blocks, or things to throw when angry. A conceptual class then merely identifies those stimulus elements, or their relation, which elicits a common response within the tenancy of a particular motivational state.

It is frequently pointed out that concept formation appears to involve insight. Once subjects have obtained a relational concept it is often transferable to a wide range of stimulus variations. In a discriminative situation where the correct response is to pick the longer of two stimuli presented, the habit, once learned, is not degraded by a change in the stimuli.

While it is particularly to the credit of the Gestalt psychologists (15,24) that attention has been focused on the phenomena of insight, they did not get far beyond stating the problem. They recognized that insight involved a rather sudden and dramatic reorganization of the forces in the psychological field, but they did not clarify the process by which this was made possible.

Despite the ingeniousness of some of the explanations, especially Hull's habit family hierarchy (11), insightful behavior continues to be an alien thorn within the flesh of reinforcement theory. The insistence of linking stimuli to responses only within the context of reinforcement and temporal contiguity requires extensions to the theory.

One extension, the mediation hypothesis (17), is perhaps as suitable as any other for discussing the facts of concept formation. A concept refers to a set of stimuli tied together through a common response. Insight occurs when one of the stimuli is forced (through reinforcement) to attach itself to a new response. When this occurs the other stimuli become similarly attached through the mediation of the old response or some vestige of it. It is for this reason that language exercises so important a role in the development of concepts. If a subject already has a tendency to say "circle" to certain objects, it is not difficult to learn a new response to the same class of objects. The response of "circle" sets up stimuli which, it turn, elicits the new response.

Hull's (10) early work on concept formation investigated the role of common stimulus elements. Subjects were shown 12 packs of cards on which Chinese characters were written. Each card was an example of a concept (a concept being distinguished by a common stimulus element) and each pack contained one card relating to each concept. The subjects were required to make differentiating verbal responses to the cards. Hull found the development of learning in this situation to be similar to discrimination learning. Hull has been criticized for calling this type of learning concept formation. Osgood (17) prefers to describe it as labeling. This criticism is not justified in view of the fact that Hull's concepts would have operated like any other concepts had he chosen to manipulate them in the more common fashion.

Smoke (21) prefers to speak in terms of common perceptual elements rather than stimulus elements. He defines a concept as a symbolic reaction to a class of stimuli displaying common perceptual organization. Although the common element in Smoke's stimuli was relational, it is evident that the perceptual organization which evolved was forced upon the subjects by nature of the response required and the reinforcement schedule followed. To this extent Smoke's study is not different from Hull's.

An interesting method of studying concept formation has been described by Reed (18, 19,20). Subjects were required to learn a new verbal response to a previously established stimulus class. Cards were shown containing four English words, one of which was an instance of a given concept. All B E P cards included the name of a vegetable; all D A X cards the name of a color; and all K U N cards the name of an animal. Reed was able to show that subjects who were told to look for general concepts (i.e., told that the nonsense syllable represented a definite class of things) were superior

to subjects who were not given this information. This illustrates that subjects are able to link a new response to a whole class of stimuli by associating several members of that class to the response.

Heidbreder (7,8,9) has defined a concept as a logical construct which is transferable from situation to situation and is communicable between persons. Using drawings on a memory drum, she had her subjects anticipate the nonsense syllables associated with the pictures. The task, in common with the one given by Reed, was attaching a nonsense syllable response to a previously learned class of stimuli. She found that concrete objects are more easily identified as a class than spatial forms are, and the latter are more easily identifiable than abstract numbers.

Heidbreder's definition of concept formation as a "logical construct" suggests something akin to a universal class distinction that exists outside of experience. One of her classes was a group of buildings. It is difficult to understand how this category is more "logical" than any other stimulus collection that has been associated through some common response mechanism. Her definition also appears to restrict the process to organisms endowed with the gift of language. This seems unfortunate. Except for its almost infinite variety and transmitability, language is not different from other response categories. Human speech is an aid to the thought process because it provides an almost inexhaustible variety of differentiating responses. It is evident that an organism, whose potential for response is limited, is also limited in the number and complexity of the concepts it can entertain. Language is more than a means of communication between individuals; in a sense it allows the person to communicate with himself through time. That is, it provides the necessary mediating responses for complex thought. Although it undoubtedly facilitates concept formation, it is not a prerequisite to it. Any response can become "symbolic" of a stimulus class.

Numerous tests of abstracting ability have been designed for use in the diagnoses of mental patients. One very useful technique has been described by Hanfmann and Kasanin (6). Wooden blocks of various shapes, lengths, colors, etc., are to be classified into four groups. It has been found that, unlike normal subjects, patients with brain damage have great difficulty making classifications involving two concepts (height and size). They have much less difficulty classifying in terms of a single concept.

In a study by Weigl (23), normal children and normal adults were compared with patients having brain lesions. It was found that although the adults could entertain a more abstract basis for classification than the children, the latter were superior to the patients.

In most studies concept formation is treated as though it were a pure mental state, or at most a stimulus act. It should be evident, however, that the tendency to abstract is not unrelated to other variables. Response requirements and reinforcement schedules, in particular, affect the evolution of a concept. A concept does not evolve simply because there exists in the mind of the experimenter a potential for classification. Rather, it comes into being because the organism is directed toward certain responses by the method of reinforcement. It the present report a series of experiments are described in which problems (concepts), reinforcement

schedules and responses are varied for normal and retarded subjects.

Variations on the Simultaneous, the Succussive and the Similarity-Difference problems have been used. Usually a double concept was involved. The subjects had to differentiate between stimulus classes and then organize this information into a response solution. The Simultaneous and Successive problems are commonly used in two choice discrimination problems. The Similarity-Difference Problem is based upon Kreezer and Dallenbach's study (14) of opposition in children.

From the previous discussion on language it would seem that some responses would prove more suitable than others in the generation of a concept. The stimulus-response relation of some problem solutions favor some responses in preference to others. In the present report two types of manual response are compared with a verbal response.

Reinforcement has two effects: it satisfies a need, and it directs subsequent responses. In this latter role it acts by confirming or contradicting the correctness of a response. The information given through reinforcement can be controlled by varying either the probability of reinforcement or the amount of reinforcement.

It is well known that subjects not only learn, but also extinguish (12) more rapidly under 100% reinforcement than under partial reinforcement. Denny (3) has proposed that the greater resistance of partial reinforcement procedures to extinction is attributable to the secondary reinforcement properties of the situation. Other results (4) indicate that the stimuli associated with non-reinforcement as well as reinforcement function to direct the response, but that their role is limited to information transmission. In some of the studies described in the latter sections, amount and probability of reinforcement are varied in accordance with the requirements of the hypotheses being tested.

Retarded subjects were not included in all of the experiments. Anyone working in the field of mental retardation will appreciate that until the experimental methods have been worked out it is often useless to bring in mentally retarded subjects. The general approach was first to exercise the variables using normal subjects, and subsequently to apply the more acceptable procedures to an evaluation of the retarded.

SECTION 2

EFFECT OF TYPE OF RESPONSE ON THE SOLUTION OF THE SIMILARITY-DIFFERENCE PROBLEM

This study was accomplished preparatory to the experiment described in Section 5. It was believed that the latter study would profit by a preliminary investigation of the stimulus and response variables exercised here. The present study compares the rate of learning of the similarity-difference relationship under two stimulus presentations (human figures as opposed to geometric figures) and three response conditions (Direct Manual, Indirect Manual, and Verbal).

The similarity-difference presentation is an excellent example of a problem requiring a relational solution. In other words, it cannot be solved in terms of the absolute properties of the stimuli. It is presented by showing the subject two stimuli simultaneously and requiring that he respond in one manner when the relevant aspects of the stimuli are identical and in another manner when they are unalike. The relationship of the relevant elements is symbolically illustrated in simplified form in Table 1.

Table 1.

Stimulus-Response Relationships in the Similarity-Difference Problem

Stimulus Variation	Stimulus 1	Stimulus 2	Correct Response
1	A	A	R1
2	A	B	R2
3	B	A	R2
4	B	B	R1

In the context of an actual problem, the subject must not only identify the relevant characteristic of the stimuli but also must learn to react to the relationship.

A study has been described by Bruner (2) in which a comparison was made of abstract and thematic materials used in a classification problem. It was found that attaining concepts with materials which are meaningful and amenable to familiar forms of grouping leads to several difficulties. First, the

subject is likely to rely upon reasonable and familiar hypotheses about the possible groupings. Thus, he may be led into a modified form of successive scanning. Secondly, it was found that thematic material will cause certain attributes to have nonrational criteriality more readily than will abstract material. The subject will tend to retain these and will formulate hypotheses utilizing them.

It would be desirable to ascertain if the same effect is obtainable in the relational problem. For this reason, two general types of stimuli have been used in the present problem. For half of the subjects, pictures of humans were used where the sex and the age of the individuals photographed were varied. For the remaining subjects, simple geometrical forms (lines which varied in direction and color) were used as the stimuli.

The solution of a discrimination problem is not independent of the nature of the response required. It has been shown (1) that a response directed toward the stimuli (approach) facilitates the Simultaneous Problem (the subject selects one of two stimuli irrespective of its position in the lateral arrangement) but inhibits the development of the Successive Problem (the subject responds in terms of a particular position in space). On the other hand, a non-approach response tends to favor the successive solution while making the simultaneous solution more difficult. This suggests that the Simultaneous Problem is solved component-wise, i.e., without reference to the relationship between the discriminanda, while the successive solution depends upon a simple relationship (configurational learning). The approach response tends to segregate the stimulus configuration into its separate elements, while the non-approach response leaves the configuration intact.

If this conclusion is valid, it could be expected that a non-approach, manual response would also facilitate the solution of the Similarity-Difference problem, since it is entirely dependent upon the relationship of the two stimuli. For this reason, Direct Manual and Indirect Manual response treatments are included in the present design.

The role of language in concept formation is not clearly understood. It does appear, however, that the difficulty in the use of a relationship is greatly reduced if the individual is given a word which is symbolic of the concept. Language facilitates learning because it acts as a mediating response. Since it permits an almost infinite variety of discriminable responses, it is more easily utilized than other response categories. It is the purpose of the present study to determine whether subjects can solve relational problems more easily using verbal responses or manual responses when the number of differentiating responses required is very limited.

METHOD

Subjects. Forty-eight normal grade school students obtained from the Fort Worth Public School System served in this study. They were randomly assigned to the twelve treatment conditions described below. The general I.Q. of the group (as determined by the Otis Test) was high. It was decided to use somewhat superior children since one of the primary purposes of this

study was to evaluate a methodology and there is less likelihood of obtaining a spurious result with homogeneous groups. All of the subjects were from the Sixth grade and were approximately within the same age level.

Apparatus. The discrimination apparatus allowed for the simultaneous exposure of two 35mm Kodachrome slides. The front panel of the apparatus (the side facing the subject) contained two projection windows for the presentation of stimulus pictures. Situated between these windows were two push-button electrical switches in vertical array. The switches were used by the subjects of the Indirect Manual Response groups. Other push-button switches were located beneath each window. These switches were used by the Direct Manual Response groups. Toward the bottom of the panel were two indicator lights. A grean light indicated that a correct response had been made, while a red light indicated a wrong response.

On the side of the console facing the experimenter, controls were available for energizing the apparatus, changing the film holders, and for programing the right-wrong indicator lights. In addition, a clock was mounted on the experimenter's console. The clock started when the stimuli were presented and terminated when a response switch was pressed. In the case where the subject's response was verbal it was necessary for the experimenter to stop the clock manually. By energizing the proper circuit, the experimenter was able to indicate the correctness of a verbal response. When manual responses were involved, however, the correct or the incorrect light would come on automatically.

Two sets of stimulus pictures were used. One set (geometrical) showed straight lines possessing a combination of the following alternatives: (1) either vertical or horizontal, (2) either red or blue, (3) either thick or thin, and (4) with either a white or a black background. The other set of stimuli showed human figures which differed in the following alternatives: (1) either male or female, (2) either child or adult, (3) either negro or white, and (4) either with the subjects sitting or standing. Consequently, each set of stimuli consisted of sixteen pictures, which varied in all possible combinations of the four differentiating aspects.

Presenting two pictures at a time (one in each of the projectors), there were 120 possible combinations of the sixteen pictures. In addition, eight trials were presented in which pictures were paired with identical pictures. Thus, the program consisted of 128 trials for both the geometrical and the humanoid stimuli. With the exception that there were not more than three identical correct responses in sequence, the order of presentation of the stimulus pairs was random.

For one of the groups presented the geometrical stimuli, the relevant aspect was whether the lines were vertically or horizontally oriented. All other variations were irrelevant. Thus, when the two pictures corresponded in orientation (whether vertical or horizontal), one response was required. However, when they were not in agreement, a separate response was required. For the other groups using the geometrical stimuli, the relevant aspect was the color of the lines. The remaining attributes were irrelevant. For half

of the subjects receiving humanoid stimuli, the sex variation was relevant. The other half of the subjects utilized the age of the person photographed as the proper basis for discrimination.

Procedure. The design of the experiment is presented in Table 2.

Table 2.

Mean Correct Responses for Groups and Treatment Conditions

	STIMULI Humanoid		Geometrical		
Relevant Stimulus Aspect	Sex	Age	Orientation	Color	Total
Direct Manual Response	M= 62.75	M= 73.75	M= 62.25	M=62.75	M=65.38
Indirect Manual Response	M= 65.50	M= 93.75	M= 77.75	M=87.50	M=81.13
Verbal Response	M=101.75	M= 91.25	M=102.00	M=78.00	M=93.25
TOTAL	M= 76.67	M= 86.25	M= 80.67	M=76.08	M=79.92

It can be seen that a 3x4 factorial design was involved which contained four subjects in each cell.

All of the subjects of the humanoid stimulus groups received the same stimulus presentations. The difference was that for one treatment, the sex of the persons photographed was the relevant feature, while for the other treatment, age was relevant. Consequently, the correct responses of the two groups would frequently differ.

Similarly, for the groups receiving the geometrical stimuli, the presentations were identical and the correct responses differed. For half of these subjects, the orientation (vertical or horizontal) of the stimuli was the significant aspect. For the other subjects the color (blue or red) of the lines was relevant to the correct solution.

The response variable had three variations. The Direct Manual treatment groups were instructed to respond by pressing one or other of the two switches which were located immediately below the display windows. For half of this group the correct response was to press the right-hand switch when the relevant aspects of the stimuli were identical, and the left-hand switch when they were different. For the other half of these subjects the relationship was reversed.

The Indirect Manual treatment groups were instructed to use the two switches which were vertically arranged with respect to one another in the center of the console. This arrangement prevented either switch from being particularly associated (due to spatial proximity) with either window. For half of the subjects the correct response was to press the upper switch when the relevant stimulus aspects were identical, and the lower switch when they were different. The other half was given the reverse solution.

The Verbal groups were required to respond by saying either "rope" or "leather" when the stimuli was presented. For half of these subjects, "rope" was correct when the relevant stimulus aspects were identical. For the other subjects, "leather" was correct when the relevant stimulus aspects were identical.

Each subject was given 128 trials under the conditions appropriate to his group. No time limit was placed on responding. The experimenter recorded both responses and response time on individual data sheets.

RESULTS

The means and the treatments for the groups are presented in Table 2. The means are based on the number of correct responses out of the 128 trials. It may be seen that the differences between stimulus treatments are small and the differences between the response variations are large. The Verbal condition was superior to the Indirect Manual condition, and the Indirect Manual condition was superior to the Direct Manual condition.

The results of the analysis of variance of the correct scores are presented in Table 3. Of the primary effects, the response variations are significant ($P < .01$) but the variance generated by stimulus difference lacks significance. The interaction of the stimulus variable with the response variable is also lacking in significance.

Table 3.

Analysis of Variance of Correct Responses

SOURCE	S.S.	df	M.S.	F	P
Stimuli	791.17	3	263.72	.61	N.S.
Responses	6251.17	2	3125.59	7.27	.01
S x R	2947.33	6	491.22	1.14	N.S.
Within	15480.00	36	430.00		
Total	25469.67	47			

Since the differences between stimulus treatments are without statistical significance, the scores may be pooled for the purpose of comparing the response treatments. The results of these tests are shown in Table 4.

Table 4.

Comparison of Response Treatments
Using Duncan's New Multiple Range Test

		D.M.	I.M.	V.	Shortest Significant Ranges
	Means	65.38	81.13	93.25	
D.M.	65.38		15.75*	27.87*	14.99
I.M.	81.13			12.12	15.75

* $p < .05$

D.M. - Direct Manual
I.M. - Indirect Manual
V. - Verbal

There exists a significant difference between the Direct Manual and the Indirect Manual treatments. The difference between the Direct Manual and the Verbal treatment is also one in which confidence can be placed. The difference between the Indirect Manual and the Verbal treatments fails to achieve significance.

As described earlier, the experimenter recorded the time to respond in each trial as well as the correctness of the response. The mean response time for each group and treatment is shown in Table 5. Again, there does not appear to be any differences of importance between stimulus treatments. The measurement favored the Indirect Manual treatment as compared to the Direct Manual. Since the Verbal treatment required the experimenter to stop the clock (with the manual groups the clock stopped automatically with the response of the subject), it was not felt that these scores should be subjected to an analysis of variance treatment. A comparison of the manual groups (ignoring stimulus treatments) by the t test failed to show significance.

Table 5.

Mean Response Time

	STIMULUS VARIABLE				
	Humanoid		Geometrical		
RESPONSE VARIABLE	Sex	Age	Vertical-Horizontal	Red-Blue	Total
Direct Manual	7.1	8.0	6.5	8.6	7.5
Indirect Manual	6.8	7.3	7.0	6.4	6.8
Verbal	7.8	7.9	7.1	7.6	7.6
Total	7.2	7.7	6.9	7.5	

DISCUSSION

It was hypothesized that the Direct Manual treatment would be inferior to the Indirect Manual procedure. This followed from the assumption that a relational problem is facilitated by a response which does not tend to individuate the stimuli into separate components. This notion was confirmed by the data.

Furthermore, it was expected that the Verbal treatment would prove superior to either of the manual treatments. This notion was supported by the data. The difference between the Direct Manual and the Verbal treatments was significant. The difference between the Indirect Manual and the Verbal treatments lacked significance, although the difference was in the predicted direction.

No differences of significance were obtained between stimulus treatments. It is difficult to interpret this result. It is not known what might have been the case had the groups been larger. It is doubtful, however, if a strong effect would have been obtained with an increase in the number of subjects. It cannot be said, of course, that the stimulus treatments were random samples from all possible stimulus treatments involving human and geometrical variations. While the geometrical stimuli were structurally more simple (less detail and interpicture variation), it may be that the relevant aspects of the humanoid stimuli (sex and age) possess more distinguishing characteristics than orientation or color in geometrical displays. Simplicity could consequently have been offset by attentional value.

SUMMARY

Forty-eight normal sixth grade students were assigned to twelve treatment conditions (4 stimulus variations x 3 response variations). All subjects were given 128 trials on the Similarity-Difference Problem. Twenty-four of the subjects saw pictures of humans which varied in sex, age, race and position. For half of these subjects, age was the relevant cue, while for the remaining subjects, sex was relevant. The other twenty-four subjects saw pictures in which straight lines were presented. These stimuli varied in orientation, color, thickness, and background. For half of these subjects, orientation was relevant, while for the other half color was relevant.

One-third of the subjects were given the Direct Manual Response condition. They were required to discriminate between two push-button switches located immediately below the stimulus windows. Another third of the subjects (Indirect Manual) discriminated using switches located between the stimulus windows. The rest of the subjects were required to respond by saying either "rope" or "leather" to each presentation. It was hypothesized that the Verbal treatment would be superior to the Indirect Manual treatment, which, in turn, would be superior to the Direct Manual treatment. On the basis of total correct responses, the prediction was upheld. Although the difference between the Verbal treatment and the Indirect Manual treatment did not achieve statistical significance, the other differences were significant.

The differences between stimulus treatments were not significant. There is no evidence, on the basis of these data, that the variations present (whether relevant or irrelevant) in the humanoid stimuli gave rise to nonrational criteriality.

This study indicates that in the development of a concept the nature of the response is of equal importance as the arrangement of the stimulus materials. Generalizing further, it can be said that in order to facilitate learning, the response required should be one which does not interfere with the process by which the concept is obtained.

SECTION 3

COMPARISON OF FOUR GRADE LEVELS ON THE SIMILARITY-DIFFERENCE PROBLEM

The purpose of the study described here was threefold: First, it was conducted to determine whether the Similarity-Difference problem differentiates between subjects on the basis of age level. Second, it was used to determine which of four stimulus aspects was the most discriminable; in connection with this aim it was also of interest to learn if an interaction exists between age level and stimulus aspect. Third, it was the purpose of this study to determine if a verbal response facilitates learning as compared to a manual response.

METHOD

Subjects. A total of 128 normal subjects were tested. These were equally divided between boys and girls. Thirty-two subjects were tested in each of four grade levels: first (G1), third (G3), fifth (G5), and seventh (G7). Socio-economically, they could be described as coming from families ranging from the lower to the middle class bracket. They were obtained from the Fort Worth schools and from surrounding school districts.

It was originally planned to use negro children for half of the subjects. However, the local tension over the school integration issue did not make this feasible. In particular, the principals of the negro schools did not want their students brought into comparison with white children.

The subjects were all within the age span that was normal for their particular grade. For example, all of the first graders were either six or seven years old, depending upon how their birth dates fell.

The subjects were volunteers. The teacher or school counselor asked them if they would care to participate. If a child was agreeable, he was given a letter addressed to his parents requesting permission to do the testing.

Apparatus and Stimuli. The apparatus consisted of a vertical panel on which the stimulus pictures could be displayed. It was painted black and was 13 inches long and 4 inches high. Set in the panel were two slots in which the pictures were placed. The picture slots were side by side, and were 3 inches apart. Arranged vertically between the picture slots were two push buttons for use by the Indirect Manual groups. The subject and the experimenter faced one another across a small table with the picture rack between. The experimenter changed the pictures from trial to trial, placing them in the slots facing the subject.

Two sets of stimuli were used. In the first phase of the experiment each stimulus was either the letter A, or the letter B. Since both picture slots were used there were four possible combinations of stimuli: AA, BB, AB, and BA. In the first phase these combinations were presented in a prearranged order of forty-eight trials. The same order was used for all subjects.

In the second phase of the experiment a different set of stimuli were used. These consisted of 32 photographs of humans who varied in terms of age (adult or child - S1), object (either shown holding a large ball or not - S2), race (white or negro - S3), sex (male or female - S4), and facial expression (smiling or not smiling - S5). The photographs were 3"x 3" black and white contact prints mounted on 4"x 5" cards. On each trial the experimenter presented two of these photographs simultaneously. The various combinations of stimulus variables are shown in Table 6.

Procedure (Phase 1)

All subjects were given 48 trials involving the Similarity-Difference principle with the AB stimuli. Half of the subjects (R2) were given the Indirect Manual treatment and were required to press one of the two push-button switches when the stimulus cards were similar (AA or BB), and the other push-button switch when the stimuli were different (AB or BA). The response procedure was counterbalanced between these subjects. Half of the R2 subjects pressed the upper switch when the stimuli were identical and the lower switch when they were different. For the remaining R2 subjects this relationship was reversed.

The remaining half of the subjects responded verbally (R1) rather than manually by saying either "X" or "Y" to the stimulus presentation. Again the correctness of the responses "X" and "Y" was counterbalanced with respect to the similarity-difference relationship. Reinforcement was given by telling the subject when his response was correct, and also by giving him one M&M chocolate candy for each correct response.

RESULTS (PHASE 1)

Two methods of evaluation were used in comparing these results: (1) trials to criterion, and (2) errors. Criterion was achieved when the subject had completed twelve successive errorless responses. The results for this method of measurement are presented in Table 7. This shows that the number of trials to criterion decreased as a function of grade level, and that there is little difference between the manual and the verbal response treatments except in the case of the seventh grade subjects. At this grade level the results definitely favor the manual response treatment.

The results of an analysis of variance of the trials to criterion is presented in Table 8. The significance of the difference between treatments and groups is shown as a matrix of $\underline{t}$ scores in Table 9.

Table 6.

Description of the Thirty-Two Stimuli Used in Phase 2

	STIMULUS				
	Age	Object	Race	Sex	Facial Expression
1	Adult	Ball	Negro	Male	Smiling
2	"	No ball	"	"	"
3	"	Ball	"	"	Non smiling
4	"	No ball	"	"	" "
5	Child	Ball	Negro	Male	Smiling
6	"	No ball	"	"	"
7	"	Ball	"	"	Non smiling
8	"	No ball	"	"	" "
9	Adult	Ball	White	Male	Smiling
10	"	No ball	"	"	"
11	"	Ball	"	"	Non smiling
12	"	No ball	"	"	" "
13	Child	Ball	White	Male	Smiling
14	"	No ball	"	"	"
15	"	Ball	"	"	Non smiling
16	"	No ball	"	"	" "
17	Adult	Ball	Negro	Female	Smiling
18	"	No ball	"	"	"
19	"	Ball	"	"	Non smiling
20	"	No ball	"	"	" "
21	Child	Ball	Negro	Female	Smiling
22	"	No ball	"	"	"
23	"	Ball	"	"	Non smiling
24	"	No ball	"	"	" "
25	Adult	Ball	White	Female	Smiling
26	"	No ball	"	"	"
27	"	Ball	"	"	Non smiling
28	"	No ball	"	"	" "
29	Child	Ball	White	Female	Smiling
30	"	No ball	"	"	"
31	"	Ball	"	"	Non smiling
32	"	No ball	"	"	" "

Table 7.

Means Based upon Groups and Treatments
of Trials to Criterion, Phase 1

GRADE	RESPONSE		TOTAL
	Verbal (R1)	Manual (R2)	
G1	45.50	44.75	45.13
G3	40.81	43.25	42.03
G5	39.06	44.25	41.66
G7	40.00	29.25	34.63
Total	41.34	40.38	

Table 8.

Analysis of Trials to Criterion in Phase 1

	Sum of Squares	df	Mean Square	F ratio	P
Response	30	1	30	1	
Grade	189	3	63	1	
Interaction	2863	3	954.33	7.74	<.01
Within Groups	14805	120	123.37		
Total	17887	127			

Table 9.

Matrix of t Scores Based upon Trials to Criterion

Grouped by Grades and Responses for Phase 1

	G1xR2	G3xR1	G3xR2	G5xR1	G5xR2	G7xR1	G7xR2
G1xR1	.252	1.329	.740	1.731	.411	1.558	3.727**
G1xR2		1.110	.475	1.490	.158	1.301	3.491**
G3xR1			.656	.408	.925	.194	2.384*
G3xR2				1.080	.308	.881	3.125**
G5xR1					1.341	.219	1.974
G5xR2						1.152	3.348**
G7xR1							2.221*

*5% level of confidence
**1% level of confidence

The differences of significance appear to be entirely attributable to the very low scores made by the 7th grade subjects, using the Manual Response Method.

The second method of evaluation used in this phase was total errors. A tabulation of the means are given in Table 10 for the grade levels and response groupings. Errors are inversely related to grade level (although the difference between the third and fifth grades is small), and in the

Table 10.

Mean Errors Based upon Groups and Treatments, Phase 1

GRADE	RESPONSE		TOTAL
	Verbal	Manual	
G1	18.88	23.94	21.41
G3	16.19	18.81	17.50
G5	14.63	20.13	17.38
G7	16.75	9.31	13.03
Total	16.61	18.05	

aggregate, more errors are made by the Manual Response groups than by the Verbal Response groups. The differences are analyzed in Table 11 where an analysis of variance is shown.

Table 11.

Analysis of Errors in Phase 1

	Sum of Squares	df	Mean Square	F Ratio	P
Response	66	1	66	1	
Grade	1124	3	374.66	4.98	$<.01$
Interaction	879	3	293.00	3.90	$<.01$
Within Groups	9015	120	75.12		
Total	11084	127			

A matrix of *t* scores between treatments and groupings are shown in Table 12. The manual response appeared to be particularly difficult for the first grade subjects, but especially easy for the seventh grade subjects.

Table 12.

Matrix of *t* Scores Based upon Errors
Grouped by Grades and Responses, Phase 1

	G1xR2	G2xR1	G3xR2	G5xR1	G5xR2	G7xR1	G7xR2
G1xR1	1.847	.961	.029	1.592	.468	.722	3.289**
G1xR2		2.437*	1.845	3.042**	1.245	2.172*	4.474**
G3xR1			.913	.497	1.255	.178	2.066*
G3xR2				1.526	.482	.896	3.220**
G5xR1					1.827	.652	1.652
G5xR2						1.040	3.360**
G7xR1							2.150*

* 5% level of confidence

** 1% level of confidence

Procedure (Phase 2)

At the termination of Phase 1, the similarity-difference principle was explained to the subject, regardless of what his score had been. It was explained using the stimulus cards of Phase 1. After the subject understood the principle, at least within the context of the AB stimuli, he was tested on Phase 2. In this second phase, those subjects who had originally responded manually were instructed to continue to do so. Those who had been given the verbal problem were enjoined to continue in this manner of response.

The stimulus presentation was the same for all subjects. The photographs described earlier which varied in race, age, sex, object, and facial expression, were presented in pairs to the subjects.

While the stimulus presentations were identical for all subjects, the solution to the problem differed between groups. For one treatment (S1), the age of the persons photographed was relevant to the solution of the Similarity-Difference problem. In other words, if the people shown in the two pictures were both adults or both children, one response was correct; however, if one of the photographs showed a child while the other showed an adult, the opposite response was correct. All other stimulus variables were irrelevant to the problem for this treatment. For a second treatment (S2), the presence or absence of the ball was the relevant feature, and the other aspects were irrelevant. The third stimulus treatment (S3) involved the race variable, while the fourth (S4) depended upon the sex identification of the persons photographed. A description of the photographs is given in Table 6.

In all, 32 treatment groups can be delineated (4 grades x two response types x four relevant stimulus solutions). Each treatment group consisted of four subjects, two of which were girls. A breakdown of the variables is given in Table 13. Within the restrictions described, assignment of subjects to treatment groups was random. Each subject was given 128 trials on this phase of the study unless his performance reached criterion before this number of trials had been given. Criterion was arbitrarily defined as twelve successive trials without errors. This procedure seemed justified since a previous pilot study revealed that when the subject finally got the "idea", there was no longer an accumulation of errors. Again the subjects were told when their response was correct and were given an M&M chocolate candy for each correct response.

RESULTS (PHASE 2)

Again, two methods of assessment were used: the number of trials to criterion, and the number of errors. The mean number of trials to criterion for the groups and for the treatment conditions is given in Table 14. Of the stimuli, racial differentiation was the easiest, followed by sex differences and presence of object discrimination. Age difference appeared to be the most difficult to distinguish.

Table 13.

Subject Assignment to Response and Stimulus Variables

GRADE	N	RESPONSE	RELEVANT STIMULI ON PHASE 2 S1 Age	S2 Object	S3 Race	S4 Sex
G1	4	R1	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
"	"	R2	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
G3	4	R1	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
"	"	R2	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
G5	4	R1	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
"	"	R2	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
G7	4	R1	x			
"	"	"		x		
"	"	"			x	
"	"	"				x
"	"	R2	x			
"	"	"		x		
"	"	"			x	
"	"	"				x

Table 14.

Means of the Groups and Treatments for Trials to Criterion, Phase 2

A - VERBAL RESPONSE	S1	S2	S3	S4	Total
G1	88.50	88.25	61.00	50.25	72.00
G3	44.50	94.50	21.75	51.50	53.06
G5	37.75	77.00	21.00	23.75	39.88
G7	77.00	65.00	60.50	68.25	67.69
Total	61.94	81.19	41.06	48.44	58.15

B - MANUAL RESPONSE	S1	S2	S3	S4	Total
G1	101.25	54.00	28.00	89.50	68.19
G3	92.25	51.50	27.25	88.50	64.88
G5	86.75	36.50	26.75	41.75	49.94
G7	108.50	46.00	56.25	50.25	65.25
Total	97.19	47.00	34.56	67.50	62.06

C - VERBAL AND MANUAL RESPONSE	S1	S2	S3	S4	Total
G1	94.87	71.12	44.50	69.87	70.09
G3	68.37	73.00	24.50	70.00	58.97
G5	62.25	56.75	23.87	32.75	43.91
G7	92.75	55.50	58.37	59.25	66.47
Total	79.56	64.09	37.81	57.97	60.10

On the whole, the fifth grade subjects were superior to the third grade, which in turn were superior to the seventh grade subjects. The first grade required the greatest number of trials to criterion. In view of the Phase 1 results (as well as the natural expectancy that the higher grades would produce better scores), the position of the seventh grade is difficult to interpret. Considering all groups simultaneously, response differences were small, although the subjects using the verbal response required a somewhat fewer number of trials to achieve criterion than did the other subjects.

The principle effects along with their interactions are analyzed in Table 15. In this analysis, the differences between stimuli are significant ($P < .01$), as are the differences between grades ($P < .01$). The difference between the response variables is not significant, but the interaction between response and stimuli is significant ($P < .01$)

Table 15.

Analysis of Variance of Trials to Criterion, Phase 2

Source of Variation	Sum of Squares	df	Mean Squares	F	P
S: Stimuli	28,664	3	9555	5.99	<.01
G: Grade	12,918	3	4306	2.70	<.05
R: Response	370	1	370		
S x G	9,906	9	1101		
S x R	22,166	3	7389	4.63	<.01
G x R	1,431	3	477		
S x G x R	7,207	9	801		
Within	153,123	96	1595		
Total	235,785	127			

Turning now to errors, Table 16 shows that the fewest errors were made when race was relevant. This was followed by sex as the relevant variable, and then object as the relevant variable. The discrimination based upon age differences is again shown to be the most difficult.

The verbal response is slightly superior to the manual response in Table 16. The differences between grades favor the fifth grade over the other grades, which do not differ greatly from one another.

Table 16.

Mean Errors of the Groups
and Treatments, Phase 2

A - VERBAL RESPONSE

	S1	S2	S3	S4	Total
G1	36.00	34.75	21.50	15.50	26.94
G3	18.75	40.50	3.00	19.25	20.38
G5	9.75	26.75	3.00	6.25	11.44
G7	22.00	28.50	19.00	28.00	24.38
Total	21.63	32.63	11.63	17.25	20.78

B - MANUAL RESPONSE

	S1	S2	S3	S4	Total
G1	46.50	15.75	8.00	32.50	25.69
G3	36.50	12.25	8.25	38.50	23.88
G5	33.75	10.25	6.25	10.75	15.25
G7	48.25	11.75	22.00	20.25	25.56
Total	41.25	12.50	11.13	25.50	22.59

C - VERBAL AND MANUAL RESPONSE

	S1	S2	S3	S4	Total
G1	41.25	25.25	14.75	24.00	26.31
G3	27.62	26.37	5.62	28.87	22.13
G5	21.75	18.50	4.62	8.50	13.34
G7	35.12	20.12	20.50	24.14	24.97
Total	31.44	22.56	11.38	21.37	21.68

An analysis of these differences is provided in Table 17. Of the principle effects, only the variance due to stimuli is significant, although the grade differences come close to achieving the 95% level of confidence. Among the interactions, the S x R (stimuli x response) is the only significant comparison ($P < .01$).

Table 17.

Analysis of Variance of Errors in Phase 2

Source of Variation	Sum of Squares	df	Mean Errors	F	P
S: Stimuli	6,473	3	2158	4.03	<.01
G: Grade	3,262	3	1087	2.17	
R: Response	104	1	104	.21	
S x G	2,115	9	235	.47	
S x R	6,763	3	2254	4.49	<.01
G x R	135	3	45	.09	
S x G x R	1,741	9	193	.38	
Within	48,228	96	502		
Total	68,821	127			

DISCUSSION

Were it not for the results of the seventh grade subjects, the data of this study would be simple to explain. If only G1, G3 and G5 are considered, the Phase 1 data would show that problem difficulty decreases with age and that the Verbal response facilitates learning (Table 10). The G7 subjects found the Manual response much easier than the Verbal response, however. This produces an interaction that cannot be explained. It is difficult to imagine that seventh graders are relatively inept in the use of verbal symbols as compared to the fifth grade. It is equally difficult to believe that something occurs in the maturation process between the fifth and seventh grades which greatly facilitates learning by the Manual response method.

The Phase 2 results are also more difficult to explain due to the data produced by the seventh grade. For some reason G7 was inferior to G3 and G5.

Comparison of the relevant stimulus aspects shows that race is most easily distinguished. Age differences were the most difficult for these children to identify.

SUMMARY

Subjects taken from four grades (first, third, fifth, and seventh) were compared on the Similarity-Difference problem. Half of the subjects responded manually while the other half responded verbally. In the course of the experiment four stimulus aspects (age, race, sex, and presence of an object) were compared for ease of discrimination.

In general, the following results were obtained:

1. Problem difficulty decreases with age.

2. Race differences are the most easily distinguished, while age differences are the most difficult to distinguish.

3. Learning is facilitated by the Verbal Response method as compared to the Manual Response method. This statement did not hold for the seventh grade subjects.

SECTION 4

CONCEPT REVERSAL UNDER TWO PROBLEM CONDITIONS

Spence (22) has proposed that there exists three levels of perceptual organization in discriminative learning; these have been called component learning, compound learning, and learning by transverse patterning.

Component learning presumably is used in the solution of the Simultaneous Problem, which is shown in a general way in Table 18.

Table 18.

Stimulus-Response Relationships in the Simultaneous, Successive and Similarity-Difference Problems

(The underlined stimuli indicate the correct choice)

PROBLEM	VARIATION	STIMULUS 1	STIMULUS 2
Simultaneous	1	$\underline{A}$	B
	2	B	$\underline{A}$
Successive	1	$\underline{A}$	A
	2	B	$\underline{B}$
Similarity-Difference	1	$\underline{A}$	A
	2	$\underline{B}$	B
	3	A	$\underline{B}$
	4	B	$\underline{A}$

Since A is reinforced and B is not reinforced, the habit of selecting A ($S_A \longrightarrow R_A$) grows, while the habit of selecting B ($S_B \longrightarrow R_B$) does not.

It is evident that this type of organization will not solve the successive problem, which is also shown in Table 18. Reinforcement of the habit $S_A \longrightarrow R_A$ is contingent upon the position of S_A. The same is true of the

habit $S_B \rightarrow R_B$. Positions (left and right) must consequently be included in the solution of the problem. $S_A + S_L \rightarrow R$ is reinforced, as is $S_B + S_R$. $S_A \rightarrow R$ and $S_B + S_L \rightarrow R$ are not reinforced. Spence calls the perceptual organization leading to this solution Compound Learning.

Reacting to the compound of stimulus and position will not solve the Similarity-Difference Problem. In this problem (see Table 18) the correctness of a choice depends upon the stimulus, its position, and the stimulus it is paired with. In Spence's vocabulary, the solution of this problem involves transverse patterning. The reinforced habit involves the positive stimulus, its position and the stimulus trace persisting from having looked at the negative stimulus.

This method of analysis says nothing about concepts. Habits are conceived to grow entirely as a function of the stimulus-response relations which are reinforced.

If this analysis is valid, then any stimulus complex which is systematically related to a response pattern within the context of reinforcement would acquire habit strength. If a situation is presented in which more than one solution (a systematic group of stimulus-response relationships) is available, it is predicted that all solutions will acquire habit strength.

As against this view, is the Lashley-Krechevsky (16) contention that subjects selectively abstract or attend to particular aspects of the total stimulus situation, and only those aspects being responded to are affected by reinforcement.

The experimental results surrounding this controversy have tended to favor the non-continuity position of Spence. These experiments have usually involved a reversal of stimulus cues rather than a reversal in concepts, however.

The hypothesis being tested in this study is that subjects respond to all stimulus aspects which can be related to the perceptual process being used. They do not, however, respond to those aspects which require a different organizational process than the one being used.

Suppose, for example, that after a subject has learned the Simultaneous Problem involving the stimuli A and B, new stimuli are introduced. These new stimuli, C and D, are presented in a conjoined manner with the old stimuli. It is predicted that if the new stimuli can be organized within the framework of the Simultaneous solution, then they will be integrated into the stimulus complex which makes up the habit. If they cannot be organized in this manner (they might be arranged as a Successive Problem), then the subject does not learn of their relationship to reinforcement.

This view falls between the Continuity and the Non-Continuity positions. It agrees with the Continuity hypothesis to the extent of stating that new stimuli are integrated into the habit. It disagrees with the Continuity

position by insisting that a habit is more than an association of a stimulus with a response. More specifically, it states that a habit is built upon a particular conceptional framework.

METHOD

Subjects

Sixth-four children and adolescents served as subjects in this study. They were obtained from the local summer Y.M.C.A. camp. Their ages ranged from 9 to 14. Only male subjects were used. One-half of the subjects were from the Tarrant County Orphans Home, and the balance were from local middle class homes. All were normal with respect to physical and mental development.

Apparatus

The discrimination apparatus consisted of a display panel having slots for the exhibition of the stimuli. The stimuli consisted of 3"x 4" photographs. These were placed in pairs by the experimenter in the slots provided. They were shown in a vertical position facing the subject. The pairs of stimuli were changed from trial to trial.

Situated between the picture slots were two push-button switches in vertical array. The subjects chose between these switches in making their responses. On a given trial the experimenter placed two stimulus pictures in the slots and the subject made his discrimination by pushing one or the other of the two push buttons. The subject was then advised concerning the correctness of his choice and was rewarded with a small piece of candy for each correct response.

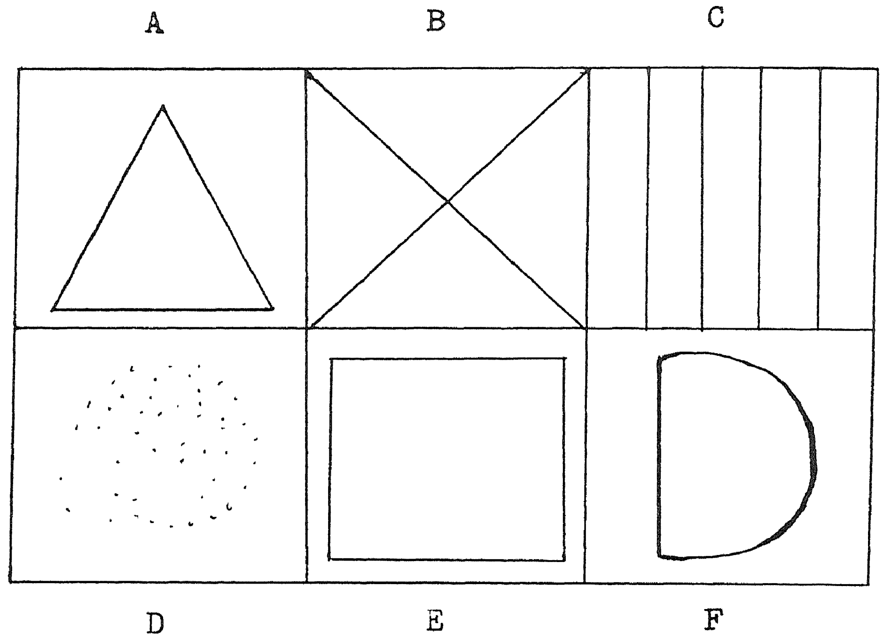
The stimuli consisted of photographs which varied in six dimensions. Each picture was divided into six equal sections. For a particular picture a section could be filled or unfilled, depending upon the stimulus variation required. In the upper left section a triangle was shown which was either resting on its base (A), or its apex (A'). In the upper center section a cross was shown with its members in either a diagonal orientation (B) or in a vertical-horizontal orientation (B'). In the upper right section four parallel lines were shown in either vertical (C) or horizontal (C') orientation. In the lower left section a group of dots were shown which varied in being either highly concentrated (D) or sparsely concentrated (D'). In the middle lower section there appeared a square which was either large (E) or small (E'). Finally, in the lower right section either the letter E(F) or the letter D(F') appeared. Examples of these variations are shown in Figure 1 with all sections filled.

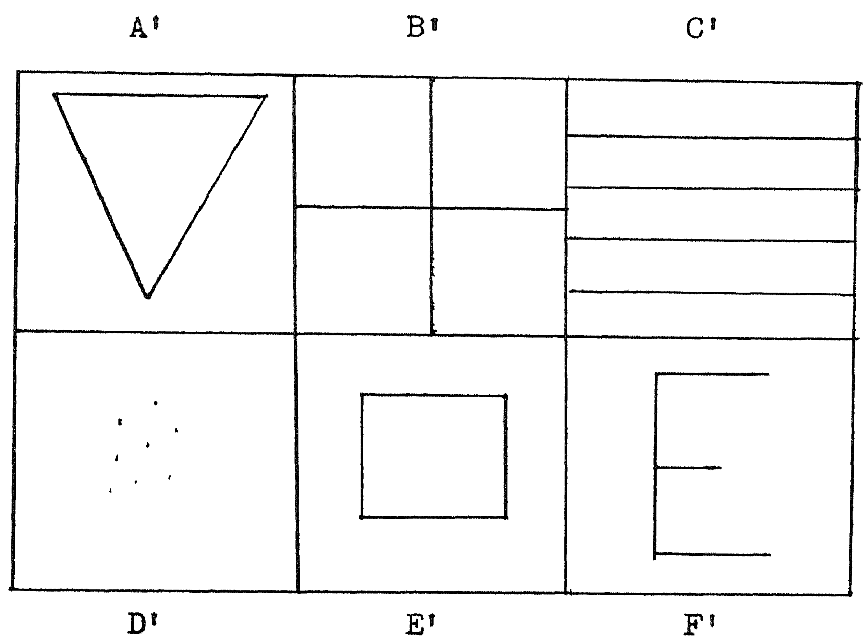
Procedure

Two basic problems were involved. These were the Similarity-Difference Problem and the Double-Successive Problem. They were selected because it was assumed that the conceptual process involved in the solution of one is

Figure 1

Example of Stimuli Showing All Sections Filled





different from that required for the solution of the other. The correct stimulus-response relationships are shown in a general way in Table 19. The designations M, M', N, N', O and O' refer to particular stimuli and their variations. For example, M could mean a large circle, while M' would designate a small circle. Similarly, N and N' would designate a separate stimulus in two discriminable variations.

It will be appreciated that the Similarity-Difference Problem is solved by making one response when the relevant stimuli are identical, and a different response when they are different. The Double-Successive Problem is solved by making one response when stimuli N or O appears, and a separate response to either of their variations. The Conjoined Problem can be solved in either of the two ways.

Table 19.

Correct Stimulus-Response Relationships of the Similarity-Difference Problem, the Double-Successive Problem and the Conjoined Problem

VARIATION		L*	R*	CORRECT RESPONSE
Similarity-Difference Problem	1	M	M	R1
	2	M'	M'	R1
	3	M	M'	R2
	4	M'	M	R2
Successive Problem	1	N	N	R1
	2	O	O	R1
	3	N'	N'	R2
	4	O'	O'	R2
Conjoined Problem	1	M N	M N	R1
	2	M'O	M'O	R1
	3	M N'	M'N'	R2
	4	M'O'	M O'	R2

* L and R refer to the position (i.e., left or right) of the stimuli in the presentation.

The experimental design of the present problem is shown in Table 20. Three stages of testing were involved. In the first stage, half of the subjects received the Similarity-Difference Problem and the remaining received the Double-Successive Problem. In the second stage a Conjoined Problem was given. For groups 1, 9, 2 and 10 a new similarity-difference

Table 20.

Experimental Design showing the Stimuli Combinations presented each group for each stage of the study.

			GROUP 1		GROUP 2		GROUP 3		GROUP 4	
STAGE	S.V.	C.R.	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	1	R_1	B	B	B	B	B	B	B	B
	2	R_1	B'	B'	B'	B'	B'	B'	B'	B'
	3	R_2	B	B'	B	B'	B	B'	B	B'
	4	R_2	B'	B	B'	B	B'	B	B'	B
2	1	R_1	ABE	ABE	ABE	ABE	ABE	ABE	ABE	ABE
	2	R_1	A'B'E'	A'B'E'	A'B'E'	A'B'E'	B'DF	B'DF	B'DF	B'DF
	3	R_1	ABE	ABE'	ABE	ABE'	ABE'	ABE'	ABE'	ABE'
	4	R_1	A'B'E'	A'B'E	A'B'E'	A'B'E	B'DF'	B'DF'	B'DF'	B'DF'
	5	R_2	ABE	A'B'E	ABE	A'B'E	A'BE	A'B'E	A'BE	A'B'E
	6	R_2	A'B'E'	ABE'	A'B'E'	ABE'	B'D'F	BD'F	B'D'F	BD'F
	7	R_2	ABE	A'B'E'	ABE	A'B'E'	A'BE'	A'B'E'	A'BE'	A'B'E'
	8	R_2	A'B'E'	ABE	A'B'E'	ABE	B'D'F'	BD'F'	B'D'F'	BD'F'
3	1	R_1	AE	AE	AE	AE	AE	AE	AE	AE
	2	R_1	A'E'	A'E'	A'E'	A'E'	DF	DF	DF	DF
	3	R_1	AE	AE'	AE	A'E	AE'	AE'	A'E	A'E
	4	R_1	A'E'	A'E	A'E'	AE'	DF'	DF'	D'T	D'T
	5	R_2	AE	A'E	AE	AE'	A'E	A'E	AE'	AE'
	6	R_2	A'E'	AE'	A'E'	A'E	D'F	D'F	DF'	DF'
	7	R_2	AE	A'E'	AE	A'E'	A'E'	A'E'	A'E'	A'E'
	8	R_2	A'E'	AE	A'E'	AE	D'F'	D'F'	D'F'	D'F'

			GROUP 5		GROUP 6		GROUP 7		GROUP 8	
STAGE	S.V.	C.R.	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	1	R_1	A	A	A	A	A	A	A	A
	2	R_1	D	D	D	D	D	D	D	D
	3	R_2	A'	A'	A'	A'	A'	A'	A'	A'
	4	R_2	D'	D'	D'	D'	D'	D'	D'	D'
2	1	R_1	ABE	ABE	ABE	ABE	ABE	ABE	ABE	ABE
	2	R_1	B'DE'	B'DE'	B'DE'	B'DE'	CDF	CDF	CDF	CDF
	3	R1	ABE	ABE'	ABE	ABE'	ABE'	ABE'	ABE'	ABE'
	4	R_1	B'DE'	B'DE	B'DE'	B'DE	CDF'	CDF'	CDF'	CDF'
	5	R_2	A'BE	A'B'E	A'BE	A'B'E	A'B'E	A'B'E	A'B'E	A'E'E
	6	R_2	B'D'S'	BD'E'	B'D'E'	BD'E'	C'D'F	C'D'F	C'D'F	C'D'F
	7	R_2	A'BE	A'B'E'	A'BE	A'B'E'	A'B'E'	A'B'E'	A'B'E'	A'B'E'
	8	R_2	B'D'E'	BD'E	B'D'E'	BD'E	C'D'F'	C'D'F'	C'D'F'	C'D'F'
3	1	R_1	BE	BE	BE	BE	BE	BE	BE	BE
	2	R_1	B'E'	B'E'	B'E'	B'E'	CF	CF	CF	CF
	3	R_1	BE	BE'	BE	B'E	BE'	BE'	B'E	B'E
	4	R_1	B'E'	B'E	B'E'	BE'	CF'	CF'	C'F	C'F
	5	R_2	BE	B'E	BE	BE'	B'E	B'E	BE'	BE'
	6	R_2	B'E'	BE'	B'E'	B'E	C'F	C'F	CF'	CF'
	7	R_2	BE	B'E'	BE	B'E'	B'E'	B'E'	B'E'	B'E'
	8	R_2	B'E'	BE	B'E'	BE	C'F'	C'F'	C'F'	C'F'

Table 20. (Cont'd)

Experimental Design showing the Stimuli Combinations presented each group for each stage of the study.

			GROUP 9		GROUP 10		GROUP 11		GROUP 12	
STAGE	S.V.	C.R.	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	1	R_1	A	A	A	A	A	A	A	A
	2	R_1	A'	A'	A'	A'	A'	A'	A'	A'
	3	R_2	A	A'	A	A'	A	A'	A	A'
	4	R_2	A'	A	A'	A	A'	A	A'	A
2	1	R_1	ABE	ABE	ABE	ABE	ABE	ABE	ABE	ABE
	2	R_1	A'B'E'	A'B'E'	A'B'E'	A'B'E'	A'CF	A'CF	A'CF	A'CF
	3	R_1	ABE	ABE'	ABE	ABE'	ABE'	ABE'	ABE'	ABE'
	4	R_1	A'B'E'	A'B'E	A'B'E'	A'B'E	A'CF'	A'CF'	A'CF'	A'CF'
	5	R_2	ABE	A'B'E	ABE	A'B'E	AB'E	A'B'E	AB'E	A'B'E
	6	R_2	A'B'E'	ABE'	A'B'E'	ABE'	A'C'F	AC'F	A'C'F	AC'F
	7	R_2	ABE	A'B'E'	ABE	A'B'E'	AB'E'	A'B'E'	AB'E'	A'B'E'
	8	R_2	A'B'E'	ABE	A'B'E'	ABE	A'C'F'	AC'F'	A'C'F'	AC'F'
3	1	R_1	BE	BE	BE	BE	BE	BE	BE	BE
	2	R_1	B'E'	B'E'	B'E'	B'E'	CF	CF	CF	CF
	3	R_1	BE	BE'	BE	B'E	BE'	BE'	B'E	B'E
	4	R_1	B'E'	B'E	B'E'	BE'	CF'	CF'	C'F	C'F
	5	R_2	BE	B'E	BE	BE'	B'E	B'E	BE'	BE'
	6	R_2	B'E'	BE'	B'E'	B'E	C'F	C'F	CF'	CF'
	7	R_2	BE	B'E'	BE	B'E'	B'E'	B'E'	B'E'	B'E'
	8	R_2	B'E'	BE	B'E'	BE	C'F'	C'F'	C'F'	C'F'
			GROUP 13		GROUP 14		GROUP 15		GROUP 16	
STAGE	S.V.	C.R.	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	1	R_1	B	B	B	B	B	B	B	B
	2	R_1	C	C	C	C	C	C	C	C
	3	R_2	B'	B'	B'	B'	B'	B'	B'	B'
	4	R_2	C'	C'	C'	C'	C'	C'	C'	C'
2	1	R_1	ABE	ABE	ABE	ABE	ABE	ABE	ABE	ABE
	2	R_1	A'CE'	A'CE'	A'CE'	A'CE'	CDF	CDF	CDF	CDF
	3	R_1	ABE	ABE'	ABE	ABE'	ABE'	ABE'	ABE'	ABE'
	4	R_1	A'CE'	A'CE	A'CE'	A'CE	CDF'	CDF'	CDF'	CDF'
	5	R_2	AB'E	A'B'E	AB'E	A'B'E	A'B'E	A'B'E	A'B'E	A'B'E
	6	R_2	A'C'E'	AC'E'	A'C'E'	AC'E'	C'D'F	C'D'F	C'D'F	C'D'F
	7	R_2	AB'E	A'B'E'	AB'E	A'B'E'	A'B'E'	A'B'E'	A'B'E'	A'B'E'
	8	R_2	A'C'E'	AC'E	A'C'E'	AC'E	C'D'F'	C'D'F'	C'D'F'	C'D'F'
3	1	R_1	AE	AE	AE	AE	AE	AE	AE	AE
	2	R_1	A'E'	A'E'	A'E'	A'E'	DF	DF	DF	DF
	3	R_1	AE	AE'	AE	A'E	AE'	AE'	A'E	A'E
	4	R_1	A'E'	A'E	A'E'	AE'	DF'	DF'	D'F	D'F
	5	R_2	AE	A'E	AE	AE'	A'E	A'E	AE'	AE'
	6	R_2	A'E'	AE'	A'E'	A'E	D'F	D'F	DF'	DF'
	7	R_2	AE	A'E'	AE	A'E'	A'E'	A'E'	A'E'	A'E'
	8	R_2	A'E'	AE	A'E'	AE	D'F'	D'F'	D'F'	D'F'

relationship was conjoined to the old problem. Groups 3, 11, 4, 12, 5, 13, 6 and 14 received the Similarity-Difference Problem conjoined with the Double-Successive Problem. The problem given in the first stage was again retained as one of the problems of the second stage. Groups 7, 15, 8 and 16 obtained two conjoined Successive problems in the second stage. One of these problems was identical with the first stage problem.

Thus in the second stage each group received a new problem-solution conjoined with the problem-solution of the first stage. It is important to observe that in the second stage irrelevant stimuli were also added to the displays. The arrangement of the irrelevant stimuli was not such as to lend itself to any logical solution of the problem.

In the third stage, all subjects were required to solve a problem requiring the same concept as was required for the problem added in the second stage. For half of the groups (1, 9, 3, 11, 5, 13, 7 and 15) the stimuli of relevance added in the second stage were also relevant in the third stage, while for the remaining groups the irrelevant stimuli of the second stage became relevant in the third stage. These relationships are shown in detail in Table 20.

There are three useful ways of distinguishing between groups as far as the analysis of the third stage results are concerned: (1) whether the problem given in the third stage was the Similarity-Difference or the Successive problem, (2) whether the problem introduced in the second stage and carried forward in the third stage was conceptionally identical with the original problem, and (3) whether the stimuli of relevance in the second stage were also relevant in the third stage.

Ignoring stimulus differences, which were included for purposes of control, there are eight treatment groups. These were formed by 2 problems (Similarity-Difference vs. Double-Successive) x 2 orders of concept presentations (original vs. reversal) x 2 orders of stimulus relevancy.

If a subject is able to integrate (as required in the second stage) only those stimulus-response relations which can be organized within the original conceptual framework, then (ignoring problem and stimulus differences) groups 1, 9, 7 and 15 should prove superior to all other groups. This follows from the fact that the problem introduced in the second stage and carried forward into the third stage required the same concept as the problem given in the first stage. The remaining groups should be inferior since either a new concept was introduced in the second stage (groups 3, 11, 5, and 13) or a reversal of stimulus relations was made to occur in the third stage (groups 2, 10, 8, and 16), or both of these events occurred (groups 4, 12, 6, and 14).

If this hypothesis is correct there should be no significant difference between groups 3, 11, 5, and 13 (considered as a whole) and groups 4, 12, 6, and 14. This follows from the fact that all of these groups received a new concept in Stage 2 and it can be reasoned that stimulus reversal is, as a consequence, immaterial.

On the other hand, if the continuity position is valid (that the new stimulus-response relationships of Stage 2 are integrated in the habit regardless of their conceptual context), then a marked difference between the groups receiving the original stimuli (19, 3, 11, 5, 13, 7 and 15) and those receiving the stimuli reversed (2, 10, 4, 12, 6, 14, 8 and 16) should result.

In the first stage the subjects received 90 trials. One hundred trials were given in the second stage, and 36 trials were given in the third stage. Except for the stimuli presented, the procedure did not change between stages. At the termination of any trial the subject was advised concerning the correctness of his response and was given a chocolate drop if his response was correct. The experimenter, in addition to setting up the stimulus pictures and reinforcing the subject, was responsible for keeping data sheets on the subjects' performance. Responses were recorded as either errors or as correct responses.

RESULTS

The mean errors for the three stages for all groups are shown in Table 21.

Table 21.

Mean Errors for the Groups
in the Three Stages of the Experiment

GROUP	STAGE 1	STAGE 2	STAGE 3
1	15.00	9.75	10.25
2	14.00	11.75	9.25
3	15.75	5.25	15.25
4	23.25	.75	17.25
5	15.75	13.00	14.25
6	9.75	13.00	14.25
7	29.25	2.25	11.75
8	32.00	2.50	8.50
9	24.00	7.25	6.50
10	26.00	11.75	7.75
11	6.75	6.25	10.25
12	15.50	1.00	12.25
13	14.50	13.50	13.25
14	21.25	21.00	18.00
15	31.50	13.25	5.50
16	15.75	17.75	12.25

The results for Stage 1 and Stage 2 have not been subjected to statistical analysis since they are not critical to the hypothesis being tested.

Table 22 shows the third stage results combined in terms of the variables of interest. In this table stimulus variations are ignored since they were exercised only for purposes of experimental control. The principle variables are (1) the type of problem solved in the third stage (problem), (2) whether the problem solved involved a new concept or was identical with the first concept learned (concept), and (3) whether the relevant stimuli had suffered reversal from the second stage (stimuli).

Table 22.

Mean Errors Based Upon Treatments

A - SIMILARITY-DIFFERENCE PROBLEM

	ORIGINAL CONCEPT	REVERSAL CONCEPT
Original Stimuli	M = 8.38 Groups 1 & 9	M = 13.75 Groups 5 & 13
Reversal Stimuli	M = 8.50 Groups 2 & 10	M = 16.12 Groups 6 & 14

B - SUCCESSIVE PROBLEM

	ORIGINAL CONCEPT	REVERSAL CONCEPT
Original Stimuli	M = 8.62 Groups 7 & 15	M = 12.75 Groups 3 & 11
Reversal Stimuli	M = 10.37 Groups 8 & 16	M = 14.75 Groups 4 & 12

C - PROBLEMS COMBINED

	ORIGINAL CONCEPT	REVERSAL CONCEPT	TOTAL
Original Stimuli	M = 8.50 Gps. 1,9,7 & 15	M = 13.25 Gps. 5,13,3 & 11	M = 10.87
Reversal Stimuli	M = 9.44 Gps. 2,10,8 & 16	M = 15.44 Gps. 6,14,4 & 12	M = 12.44
TOTAL	M = 8.97	M = 14.34	M = 11.65

As predicted, those subjects receiving both the original concept and the original stimuli produced superior scores compared to the other groupings. The magnitude of the difference between the groups receiving the reversal concept with the original stimuli and the groups receiving the reversal concept with the reversal stimuli is disturbing, however. The hypothesis being tested predicted no difference between these groupings.

In Table 23 the principle effects along with their interactions are analyzed. The difference between problem difficulty (Similarity-Difference vs. Double-Successive) is negligible. The difference between groups receiving a new concept (having a problem introduced in the second stage conceptually different from the one given in the first stage) and those receiving the old concept (having a problem introduced in the second stage conceptually identical with the first stage problem) is highly significant. Surprisingly, the difference between groups receiving the reversal stimulus procedure and those receiving the non-reversal stimulus procedure is not significant. All of the interactions lack statistical significance.

Table 23.

Analysis of Variance of Errors made in Stage 3

SOURCE OF VARIATION	SUM OF SQUARES	df	MEAN SQUARES	F	P
A (Problem)	.07	1	.07	-	NS
B (Concept)	462.25	1	462.25	20.65	<.01
C (Stimulus)	39.07	1	39.07	1.75	NS
A x B	20.25	1	20.25	.90	NS
A x C	1.55	1	1.55	.07	NS
B x C	6.25	1	6.25	.28	NS
A x B x C	4.00	1	4.00	.18	NS
Within	1253.00	56	22.38		
Total	1786.44	63			

DISCUSSION

These results do not lend themselves to easy interpretation. They do not particularly support the hypothesis that subjects are able to integrate only those stimulus-response relations which can be organized within the initial concept. It is true that the treatment involving concept reversal produces inferior scores compared to non-reversal treatment, but this fact can be better explained by a notion such as "set" (i.e., the predisposition of a subject to solve new problems in the same way old problems were solved). In particular, what is lacking in evidence to substantiate the hypothesis being tested is a significant interaction between concept and stimuli (B x C in the analysis). Following the logic of this hypothesis leads to the expectation of a considerable difference in the effect of stimulus reversal on the groups receiving the original concept and the groups receiving a reversal concept. The data suggests that stimulus reversal affects concept reversal and non-reversal in a non-differential fashion. This is more or less the prediction generated by the continuity theory. Since the analysis of variance did not show stimulus reversal treatment to be significantly different from the non-reversal treatment, these results do not absolutely support the continuity hypothesis. The difference is in the direction predicted by this hypothesis, however.

It is possible, of course, that the effect of "set" was strong enough to mask out the other effects sought in this experiment. Whatever the cause, the assumption that the continuity theory holds for stimulus change, but not for a change in concept, is not sustained by the results of this experiment.

SUMMARY

The hypothesis was tested that subjects integrate new stimuli into an established habit only when these stimuli can be organized within the conceptual framework already in use. The results were not in support of this hypothesis.

SECTION 5

COMPARISON OF NORMAL AND RETARDED CHILDREN ON THE SIMILARITY-DIFFERENCE PROBLEM

The experiments described in Sections 2, 3, and 4 can be considered as preliminary to the present study. In these experiments, an effort was made to find the conditions which would be sensitive to the hypotheses being presently tested. When conditions were found which would successfully differentiate between normal subjects, they were to be applied against retarded subjects as well.

The effects of several stimulus, response, and reinforcement variables are investigated in this study. The effort was not simply to demonstrate that retarded subjects are inferior to normal subjects; the variables exercised were selected in the hope of finding an interaction between them and the subject sources. To find retarded subjects inferior to normal subjects is not surprising, but to find an interaction between problem variables and subject variables can give a cue concerning the fundamental nature of retardation.

Except for the reinforcement variable, the procedures were similar to those given in the preceding studies. This variable was included to test the hypothesis that the effect of reinforcement does not vary so much as a function of its absolute value as it does to its value relative to the subject's level of expectation. It is assumed that during testing the subject adapts to a particular level of reinforcement. When the level of reinforcement (regardless of its value) decreases, instability results. The subject reacts by seeking a new solution to the problem which will bring the rate of reinforcement back to its previous level. On the other hand, if the reinforcement rate obtained is at the expected level the subject will not seek a better solution since he is satisfied with the prevailing rate.

METHOD

Subjects: Two hundred and sixteen children served in this study. Seventy-two were junior high school students (JH). Since these subjects were selected from the "advanced" classes of a school situated in an upper middle class neighborhood, it is likely that their average intelligence was somewhat higher than is usually encountered at this level.

The second group (GS) of 72 subjects was made up of fourth and fifth graders obtained from a grade school situated in a middle class to upper middle class neighorhood.

The third group (R) tested also consisted of 72 subjects. For purposes of exposition, this group is referred to as R (retarded). This appellation is not altogether accurate, however, since the majority of these subjects can be considered marginal. As such, they included both the dull-normal and

the mentally retarded. All were educable, however. Some selection was exercised with respect to this group in terms of the willingness of the subjects to cooperate and ability to sustain interest. Several of these children had to be replaced by other subjects when they persisted in stereotyped or otherwise inappropriate responses, or when it became evident that they were making no attempt to solve the problem. It is likely that few of the subjects of this group had I.Q.'s in excess of 80 or below 50. Actually, I.Q. evaluations were available on most of the subjects of the experiment, but these had been obtained in a variety of ways, making direct comparisons impractical. The ages of the R Group ran from 10 to 17 years of age.

An attempt was made to control for sex differences in the subject-treatment groupings. This was successful except in the case of the R Group, which consisted of 42 boys and 30 girls.

Apparatus: The discrimination apparatus used in this study had two 4"x 4" translucent windows on which the stimuli were projected. Behind each window was a small slide projector which handled 35mm Kodachrome transparencies. The experimenter changed these slides from trial to trial, and the subject responded differentially, depending upon which pictures were shown in the two windows. One experimental treatment (DM) required the subject to select between two electric push-button switches; one located directly below each of the two stimulus windows. A second experimental condition (IM) required the subjects to respond on each trial by pressing one of two push-button switches located, in vertical array, midway between the two stimulus windows. In a third treatment condition (V), the response requirement was verbal rather than manual and none of these switches were manipulated.

The apparatus was programed in such a manner that when a response was correct for the Direct Manual (DM) and the Indirect Manual (IM) treatments, a green indicator light came on. When the response was in error, a red indicator light came on. The words "right" and "wrong" appeared beneath the green and red lights respectively. For the Verbal Response (V) treatment, it was necessary for the experimenter to illuminate the appropriate light. This was done using switches located on the experimenter's panel.

The stimuli, which appeared in pairs in the translucent windows, consisted of 35mm color transparencies. These pictures varied in four respects: (1) age of the person photographed (child or adult), (2) sex of the person photographed (male or female), (3) race of the person photographed (negro or white), and (4) color of background (red or blue). These variations generated 16 separate photographs. A description of these variations is shown in Table 24.

Table 24

Variations in the Stimuli used in the Experiment

Variation	Age	Sex	Race	Background
1	Child	Male	Negro	Red
2	"	"	White	"
3	"	Female	Negro	"
4	"	"	White	"
5	Adult	Male	Negro	"
6	"	"	White	"
7	"	Female	Negro	"
8	"	"	White	"
9	Child	Male	Negro	Blue
10	"	"	White	"
11	"	Female	Negro	"
12	"	"	White	"
13	Adult	Male	Negro	"
14	"	"	White	"
15	"	Female	Negro	"
16	"	"	White	"

Procedure: The Similarity-Difference Problem was given to all subjects in all stages of the experiment. For the first stage of the study, 54 subject-treatment groups can be delineated. Each of these groups contained 4 subjects. These were generated by the following variables:

A - Subject Source:
1. Junior High School (JH)
2. Grade School (GS)
3. Retarded (R)

B - Stimulus of Relevance:
1. Race (NW)
2. Sex (MF)
3. Age (CA)

C - Response Category:
1. Direct Manual (DM)
2. Indirect Manual (IM)
3. Verbal (V)

D - Number of Stimulus Variations Shown:
1. Race, Sex, and Age (3V)
2. Race, Sex, Age, and Background (4V)

The stimulus of Relevance simply means that for one treatment the negro-white difference was relevant while the other differences (male-female, child-adult, and blue-red background) were irrelevant. In the second treatment, the male-female difference was relevant and the other differences were irrelevant. For the third treatment the age of the persons photographed was the relevant aspect.

Under Response Category there were three treatments: 1, the subject was instructed to respond by pressing the switches located directly below the stimulus windows (DM); 2, the subjects pressed the switches located between the windows (IM); and 3, the subjects were instructed to respond by saying either "X" or "Y" to each presentation (V). Counterbalancing between responses (R_1 or R_2) and stimuli (similar or different depending on the relevant aspect) was effected for all response treatments.

The number of stimulus variations shown differed between groups. For some subjects (4V) all four (race, age, sex, and background) differences were experienced, while for the remaining subjects (3V) background was held constant (either red or blue).

One hundred and twenty trials were given each subject in the first stage of the experiment. For the 4V treatment the presentations were a random sample (without replacing) of the 256 usable combinations of pictures. The only restriction employed was that the number of differentiating responses (R_1 and R_2) were to be equal. In other words, the same number of similar and different presentations (with respect to the stimulus variation of relevance) were made. For the 3V condition, all of the 54 available combinations were used at least once. The remaining 56 combinations were randomly drawn (without replacing) from the original 64. The order of presentation of the combinations was random.

In addition to seeing the red or green indicator light at the end of each trial, the subject was given a candy reward for each correct response in the first stage of the experiment.

The second stage of the experiment was identical with the first stage except that the indicator lights were turned off and the reinforcement rate with the candy dropped to 50%. Testing was actually continuous with the first stage. The subjects were simply told that the lights would no longer come on and candy would be given "about half of the time" for correct responses, but would not be given for incorrect responses. A total of 40 trials were given during this stage.

In the third stage the Stimulus of Relevance was changed for all subjects without their knowledge. For example, of a group (N=4) having the race concept, two subjects were put on a problem in which the age difference was relevant, and the remaining two were placed on the male-female concept. At this point it should be explained that if a subject continued to respond in terms of the original stimulus relationship, he would still be correct 50% of the time on the new relationship. This is due to the conjunction of the stimulus elements. If he had been given to expect about 50% reinforcement and he gets 50% reinforcement, it is hypothesized that he will not reorganize his approach to the problem. He will not search for a new solution. On the other hand, if he expects about 50% reinforcement and only obtains 25% reinforcement, it is hypothesized that he will recognize that some change has been effected without his knowledge and will search for a new solution to the problem.

This hypothesis was tested by raising one-half of the subjects back to 100% reinforcement on the new concept (again without their knowledge) and leaving the remaining subjects on 50% reinforcement. It can be appreciated

that those raised to 100% would continue to get 50% reinforcement by responding to the incorrect stimuli, while those left at 50% would drop to 25% if they continued in the original solution. Of the sub-group (N=4) described earlier (which were originally given the race discrimination and in the third stage were separated into groups of two, one having the age concept and the other having the sex concept), a further subdivision was made in which one subject was given 100% reinforcement in the third stage, while the other continued to receive 50%. Thus, in the experiment as a whole, there are 216 subject-treatment groups having only one subject per group.

One hundred and twenty trials were again given in the third stage of the experiment. For all stages (except the second stage), data sheets were kept on the subjects by the experimenter. On these sheets, an account of correct and incorrect responses was recorded.

RESULTS

Stage 1 - A summary of the mean correct responses for treatments is provided in Table 25.

Table 25

Summary of the Mean Correct Responses for Treatments, Stage 1

Variation	Treatment	Mean
Subject Source	JH	87.49
	GS	85.46
	R	74.94
Stimulus of Relevance	NW	85.60
	MF	82.04
	CA	80.25
Response Category	DM	79.08
	IM	82.15
	V	86.65
Number of Stimulus Variations	3V	85.19
	4V	80.06

Between subject sources the difference usually favored the junior high groups, although under certain circumstances the grade school groups were superior. With very few exceptions the grade school groups were superior to the retarded groups.

Race proved easier to discriminate than sex; and the latter, in turn, was somewhat easier than age, although there are many group comparisons where this statement does not apply.

Response Category exercises a differentiating effect. In general, the Verbal procedure is superior to the Indirect Manual procedure, and both of these are superior to the Direct Manual procedure.

As expected, the four-stimulus variable condition proved more difficult than the three-stimulus variable condition.

Due to some anomalous results in the third stage of the experiment (which will be explained later), it was decided to divide each stage into two separate analyses; one analysis is of the results of the Three-Stimulus Problem, and the second analysis takes up only the data of the Four-Stimulus Problem.

Since, however, it would be of interest to know whether an interaction exists between subject source and number of stimulus variations, this analysis is undertaken in Table 26. It will be seen that the interaction is not significant, although both primary effects achieve statistical significance.

Table 26

Analysis of Variance
Comparing Subject Sources (JH, GS, and R)
and Number of Stimulus Variations (3V and 4V), Stage 1

Source of Variation	SS	df	MS	F	P
S (subject source)	6526.73	2	3263.37	34.82	<.01
N (number of variations)	1420.91	1	1420.91	15.16	<.01
S x N	453.79	2	226.90	2.42	NS
Within	19680.94	210	93.72		
Total	28082.37	215			

In Table 27, an analysis of variance based upon Subject Source, Relevant Stimulus, and Response Category is shown for the 3V Problem. All principle effects are significant but none of the interactions achieve the 95% level of confidence.

Table 27

Analysis of Variance of the 3V Data Comparing Subject Sources (JH, GS, and R), Relevant Stimulus (NW, MF, and CA), and Response (DM, IM, and V), Stage 1

Source of Variation	SS	df	MS	F	P
A (subject source)	5133.17	2	2566.59	36.73	<.01
B (stimuli)	948.50	2	474.25	6.79	<.01
C (response)	1675.50	2	837.75	11.99	<.01
A x B	521.83	4	130.46	1.87	NS
A x C	217.33	4	54.33	.78	NS
B x C	217.17	4	54.29	.78	NS
A x B x C	670.17	8	83.77	1.20	NS
Within	5659.25	81	69.87		
Total	15042.92	107			

An analysis of the significance of sub-group differences is provided by Table 28. The differences above the heavy stairstep line are significant at the 95% level of confidence as determined by Duncan's New Multiple Range Test. Most of the differences are attributable to Subject Source.

In Table 29 an analysis is accomplished on the 4V treatment. As shown, the Subject Source and Response Category effects are significant but the difference based upon Relevant Stimuli is not significant. The interactions are similarly without significance. The Multiple Range Test is applied to these data in Table 30. Again the outstanding effect appears to be due to the subject source differences.

Stage 3 - A summary of the mean correct responses for treatments in Stage 3 is given in Table 31.

A distressing anomaly becomes evident at this juncture. The Junior High Group is considerably inferior to the Grade School Group. Although this fact in itself is not particularly disconcerting, the results of the analysis of variance shown in Table 32 are. In this table, the interaction between Subject Source and Number of Stimuli is highly significant. Checking back into the data, it was found that the Junior High Group was superior to the other groups on the Four-Stimulus Problem, but was inferior on the Three-Stimulus Problem. The small mean for this group on this problem ($M = 77.17$) brings these particular data into question. In attempting to explain this discrepancy a number of factors have been considered. While the reason cannot be identified with precision (schedule, instructions to subjects, etc., do not appear to have been at fault), it seems likely that for the JH subjects using the 3V stimuli the order of stimulus presentation became inadvertantly disarranged. This is

Table 28

Comparison of the 3V Groups by Duncan's Multiple Range Test, Stage 1

		$A_1B_3C_1$	$A_3B_1C_2$	$A_2B_2C_1$	$A_1B_3C_3$	$A_2B_2C_2$	$A_1B_2C_1$	$A_2B_3C_2$	$A_3B_1C_3$	$A_1B_3C_2$	$A_1B_1C_1$	$A_1B_2C_3$
	Means	81.75	83.00	83.75	83.75	86.25	86.50	87.25	88.50	91.00	92.75	93.25
$A_3B_2C_2$	69.25	12.50	13.75	14.50	14.50	17.00	17.25	18.00	19.25	21.75	23.50	24.00
$A_3B_2C_1$	69.75	12.00	13.25	14.00	14.00	16.50	16.75	17.50	18.75	21.25	23.00	23.50
$A_3B_1C_1$	72.50	9.25	10.50	11.25	11.25	13.75	14.00	14.75	16.00	18.50	20.25	20.75
$A_3B_3C_1$	72.50	9.25	10.50	11.25	11.25	13.75	14.00	14.75	16.00	18.50	20.25	20.75
$A_3B_3C_3$	74.50	7.25	8.50	9.25	9.25	11.75	12.00	12.75	14.00	16.50	18.25	18.75
$A_3B_2C_3$	75.00	6.75	8.00	8.75	8.75	11.25	11.50	12.25	13.50	16.00	17.75	18.25
$A_3B_3C_2$	76.00	5.75	7.00	7.75	7.75	10.25	10.50	11.25	12.50	15.00	16.75	17.25
$A_2B_3C_1$	78.75	3.00	4.25	5.00	5.00	7.50	7.75	8.50	9.75	12.25	14.00	14.50
$A_2B_1C_1$	79.75	2.00	3.25	4.00	4.00	6.50	6.75	7.50	8.75	11.25	13.00	13.50
$A_1B_3C_1$	81.75		1.25	2.00	2.00	4.50	4.75	5.50	6.75	9.25	11.00	11.50
$A_3B_1C_2$	83.00			.75	.75	3.25	3.50	4.25	5.50	8.00	9.75	10.25
$A_2B_2C_1$	83.75				0	2.50	2.75	3.50	4.75	7.25	9.00	9.50
$A_1B_3C_3$	83.75					2.50	2.75	3.50	4.75	7.25	9.00	9.50
$A_2B_2C_2$	86.25						.25	1.00	2.25	4.75	6.50	7.00
$A_1B_2C_1$	86.50							.75	2.00	4.50	6.25	6.75
$A_2B_3C_2$	87.25								1.25	3.75	5.50	6.00
$A_3B_1C_3$	88.50									2.50	4.25	4.75
$A_1B_3C_2$	91.00										1.75	2.25
$A_1B_1C_1$	92.75											.50
$A_1B_2C_3$	93.25											
$A_2B_3C_3$	93.50											
$A_1B_1C_2$	93.50											
$A_2B_1C_3$	93.75											
$A_2B_2C_3$	95.25											
$A_2B_1C_2$	95.25											
$A_1B_2C_2$	99.50											

A_1 - JH B_1 - NW C_1 - DM
A_2 - GS B_2 - MF C_2 - IM
A_3 - R B_3 - CA C_3 - V

* 95% level of confidence

Table 28 (Cont'd)

	Means	$A_2B_3C_3$	$A_1B_1C_2$	$A_2B_1C_3$	$A_2B_2C_3$	$A_2B_1C_2$	$A_1B_2C_2$	$A_1B_1C_3$	Shortest Significant Ranges*
		93.50	93.50	93.75	95.25	95.25	99.50	103.75	
$A_3B_2C_2$	69.25	24.25	24.25	24.50	26.00	26.00	30.25	34.50	R_2 = 11.83
$A_3B_2C_1$	69.75	23.75	23.75	24.00	25.50	25.50	29.75	34.00	R_3 = 12.44
$A_3B_1C_1$	72.50	21.00	21.00	21.25	22.75	22.75	27.00	31.25	R_4 = 12.85
$A_3B_3C_1$	72.50	21.00	21.00	21.25	22.75	22.75	27.00	31.25	R_5 = 13.14
$A_3B_3C_3$	74.50	19.00	19.00	19.25	20.75	20.75	25.00	29.25	R_6 = 13.37
$A_3B_2C_3$	75.00	18.50	18.50	18.75	20.25	20.25	24.50	28.75	R_7 = 13.55
$A_3B_3C_2$	76.00	17.50	17.50	17.75	19.25	19.25	23.50	27.75	R_8 = 13.70
$A_2B_3C_1$	78.75	14.75	14.75	15.00	16.50	16.50	20.75	25.00	R_9 = 13.82
$A_2B_1C_1$	79.75	13.75	13.75	14.00	15.50	15.50	19.75	24.00	R_{10}= 13.93
$A_1B_3C_1$	81.75	11.75	11.75	12.00	13.50	13.50	17.75	22.00	R_{11}= 14.02
$A_3B_1C_2$	83.00	10.50	10.50	10.75	12.25	12.25	16.50	20.75	R_{12}= 14.10
$A_2B_2C_1$	83.75	9.75	9.75	10.00	11.50	11.50	15.75	20.00	R_{13}= 14.17
$A_1B_3C_3$	83.75	9.75	9.75	10.00	11.50	11.50	15.75	20.00	R_{14}= 14.24
$A_2B_2C_2$	86.25	7.25	7.25	7.50	9.00	9.00	13.25	17.50	R_{15}= 14.29
$A_1B_2C_1$	86.50	7.00	7.00	7.25	8.75	8.75	13.00	17.25	R_{16}= 14.34
$A_2B_3C_2$	87.25	6.25	6.25	6.50	8.00	8.00	12.25	16.50	R_{17}= 14.39
$A_3B_1C_3$	88.50	5.00	5.00	5.25	6.75	6.75	11.00	15.25	R_{18}= 14.43
$A_1B_3C_2$	91.00	2.50	2.50	2.75	4.25	4.25	8.50	12.75	R_{19}= 14.46
$A_1B_1C_1$	92.75	.75	.75	1.00	2.50	2.50	6.75	11.00	R_{20}= 14.50
$A_1B_2C_3$	93.25	.25	.25	.50	2.00	2.00	6.25	10.50	R_{21}= 14.54
$A_2B_3C_3$	93.50		0	.25	1.75	1.75	6.00	10.25	R_{22}= 14.58
$A_1B_1C_2$	93.50			.25	1.75	1.75	6.00	10.25	R_{23}= 14.61
$A_2B_1C_3$	93.75				1.50	1.50	5.75	10.00	R_{24}= 14.64
$A_2B_2C_3$	95.25					0	4.25	8.50	R_{25}= 14.68
$A_2B_1C_2$	95.25						4.25	8.50	R_{26}= 14.71
$A_1B_2C_2$	99.50							4.25	R_{27}= 14.73

Table 29

Analysis of Variance of the 4V Data Comparing
Subject Sources (JH, GS, and R), Relevant Stimuli (NW, MF, and CA)
and Response Category (DM, IM, and V)

Source of Variation	SS	df	MS	F	P
A (subject source)	1847.35	2	923.68	11.39	.01
B (relevant stimulus)	238.35	2	119.18	1.47	NS
C (response)	971.80	2	485.90	5.99	.01
A x B	591.77	4	147.94	1.82	NS
A x C	298.98	4	74.75	.92	NS
B x C	42.98	4	10.75	.13	NS
A x B x C	1057.57	8	132.20	1.63	NS
Within	6569.75	81	81.11		
Total	11618.55	107			

Table 30

Comparison of the 4V Group
by Duncan's Multiple Range Test, Stage 1

		A_2C_1	A_1C_2	A_1C_3	A_2C_3	Shortest Significant Ranges*
	Means	81.25	81.67	87.67	89.33	
A_3C_2	73.25	8.00	8.42	14.42	16.08	R_2 = 7.36
A_3C_1	73.58	7.67	8.09	14.09	15.75	R_3 = 7.74
A_3C_3	75.83	5.42	5.84	11.84	13.50	R_4 = 7.99
A_2C_2	77.67	3.58	4.00	10.00	11.66	R_5 = 8.17
A_1C_1	80.33	.92	1.34	7.34	9.00	R_6 = 8.31
A_2C_1	81.25		.42	6.42	8.08	R_7 = 8.43
A_1C_2	81.67			6.00	7.66	R_8 = 8.52
A_1C_3	87.67				1.66	R_9 = 8.60

A_1 - JH C_1 - DM
A_2 - GS C_2 - IM
A_3 - R C_3 - V

* 95% level of confidence

Table 31

Summary of the Mean Correct Responses
for Treatments, Stage 3

Treatment	Variations	Mean
Subject Source	JH	84.85
	GS	91.92
	R	79.26
Stimulus of Relevance	NW	87.99
	MF	83.44
	CA	86.86
Response Category	DM	82.63
	IM	86.31
	V	87.08
Reinforcement	100%	85.40
	50%	85.29
Number of Stimulus Variations	3V	86.42
	4V	84.27

Table 32

Analysis of Variance comparing Subject Sources (JH, GS, and R)
and number of Stimulus Variations (3V and 4V), Stage 3

Source of Variation	SS	df	MS	F	P
S (Subject Source	5789.84	2	2894.92	32.46	.01
N (Number of Variations)	249.19	1	249.19	2.79	NS
S x N	4737.90	2	2368.95	26.56	.01
Within	18729.72	210	89.19		
Total	29506.65	215			

quite possible since the tester could not see the projected stimuli, and in any case there were too many variations for a tester to follow a particular subject's problem. By the time the anomaly was noticed it was too late to retest on the problem.

Whatever the cause, it seemed reasonable to throw out the data for these 36 subjects on the Third Stage of the experiment.

As a consequence, the Third Stage analysis of the 3V problem contains only GS and R subjects, while the 4V problem analysis contains all subjects.

An analysis of variance for the 3V Problem is given in Table 33. Only the difference between Subject Sources is significant. Stimulus and response differences are high but fail to achieve the 95% level of confidence. The only conclusion we can draw from this table is that the grade school subjects were superior to the retarded children.

Turning now to the 4V Problem the results are analyzed in Table 34. Again we see that for this stage, only Subject Source differences are significant. The variance due to stimulus differences is high, as is the interaction between response and reinforcement, but these also fail significance.

A Multiple Range Test is provided for the differences between Subject Sources in Table 35. The Retarded Group was significantly inferior to either of the normal groups, but the superiority of the Junior High Group over the Grade School Group fails to attain the 95% level of confidence.

Table 33

Analysis of the 3V Problem comparing
Subject Sources (GS and R only), Relevant Stimulus (NW, MF, and CA), Response (DM, IM, and V), and Reinforcement (100% and 50%), Stage 3

Source of Variation	SS	df	MS	F	P
A (Subject)	3267.01	1	3267.01	44.97	$<.01$
B (Stimulus)	466.36	2	233.18	3.21	NS
C (Response)	344.11	2	172.06	2.37	NS
D (Reinforcement)	13.34	1	13.34	.18	NS
A x B	16.86	2	8.43	.12	NS
A x C	64.11	2	32.06	.44	NS
A x D	21.13	1	21.13	.29	NS
B x C	480.31	4	120.08	1.65	NS
B x D	50.20	2	25.10	.35	NS
C x D	122.12	2	61.06	.84	NS
A x B x C	204.14	4	51.04	.70	NS
A x B x D	65.58	2	32.79	.45	NS
A x C x D	380.33	2	190.17	2.62	NS
B x C x D	226.96	4	56.74	.78	NS
A x B x C x D	206.59	4	51.65	.71	NS
Within	2615.50	36	72.65		
Total	8544.65	71			

Table 34

Analysis of the 4V Problem comparing Subject Sources (JH, GS, and R), Relevant Stimulus (NW, MF, and CA), Response (DM, IM, and V), and Reinforcement (100% and 50%), Stage 3

Source of Variation	SS	df	MS	F	P
A (Subject Source)	4537.17	2	2268.59	24.41	.01
B (Stimuli)	491.17	2	245.59	2.64	NS
C (Response)	125.17	2	62.59	.67	NS
D (Reinforcement)	57.79	1	57.79	.62	NS
A x B	177.33	4	44.33	.48	NS
A x C	231.99	4	58.00	.62	NS
A x D	134.57	2	67.29	.72	NS
B x C	81.83	4	20.46	.22	NS
B x D	91.90	2	45.95	.49	NS
C x D	449.12	2	224.56	2.42	NS
A x B x C	491.34	8	61.42	.66	NS
A x B x D	252.49	4	63.12	.68	NS
A x C x D	531.27	4	132.82	1.43	NS
B x C x D	54.44	4	13.61	.15	NS
A x B x C x D	549.17	8	68.65	.74	NS
Within	5017.50	54	92.92		
Total	13274.25	107			

Table 35

Comparison of the 4V Data by Multiple Range Test, Stage 3

		R	GS	JH	Shortest Significant Ranges
	Means	77.44	89.28	92.53	
R	77.44		11.84	15.09	R_2 = 4.55
GS	89.28			3.25	R_3 = 4.79

DISCUSSION

On the whole, the differences of significance between Subject Sources is attributable to the inferiority of the retarded subjects and not to any particular superiority of the Junior High over the Grade School Group. Some exceptions to this statement are found in Table 28.

The First Stage of the experiment was generally more effective than the Third Stage in distinguishing between the response and stimulus variations. In the aggregate, the Verbal Response Method seems somewhat superior to the Indirect Response Method, and this latter method is greatly superior to the Direct Response Method. These results are in close agreement with the results of Section 2. They are both in agreement and disagreement with the results of Section 3, where it will be remembered an interaction occurred between Grade Level and Response Method. It is difficult to reconcile the difference in the two studies. For the present, however, we are unable to distinguish between the Indirect Manual Method and the Verbal Response Method.

In agreement with the results of Section 3, race appears to be the most discriminable stimulus difference among the stimulus aspects tested. This may be true, however, only for the section of the country from which the subjects were taken. It may be that the subjects are not more "aware" of race differences than they are of sex and age differences, but that the racial dichotomy has an emotional content greater than the other variables.

As anticipated, the 3V Problem proved easier to solve than the 4V Problem. The inclusion of the additional irrelevant relation did not interact with Subject Source, however, as it was expected to.

The primary purpose of the Third Stage Problems was to determine the difference due to the reinforcement schedules. It was expected that the 50% reinforcement schedule would prove superior to the 100% rate. This prediction occurred on the 4V Problem, but a reversal of this relationship was obtained on the 3V Problem. In neither case were the differences significant, however. The outcome is still of interest, however, since in Reinforcement Theory a difference in favor of the 100% reinforcement treatment would be expected.

Most disappointing in these analyses is the absence of any interaction of treatments with Subject Sources. It is through such interactions that we can increase our understanding of mental retardation.

SUMMARY

Normal and mentally retarded subjects were compared on the Similarity-Difference Problem under a variety of stimulus, response, and reinforcement conditions. Aside from their basic inferiority the retarded subjects reacted to the experimental variables in much the same way as the normal subjects did.

In common with the previous experiments, it was found that the Direct Manual Response condition inhibits the learning of the Similarity-Difference Problem, but the Verbal Response condition facilitates learning.

Both the retarded and the normal children found the discrimination of race easier to make than the discriminations of sex or age.

SECTION 6

COMPARISON OF TWO REINFORCEMENT SCHEDULES ON PROBLEM REVERSAL

In the preceding section, the hypothesis was tested that the efficacy of reinforcement is a function of its value relative to the subject's expectation rather than its absolute value. In order to test the theory, subjects were adapted to a 50% reinforcement schedule on one problem and were then shifted to a second problem in which the stimuli remained the same but the formerly relevant cues became irrelevant and the formerly irrelevant cues became relevant. Concurrent with the shift in cue relevance, half of the subjects were put on 100% reinforcement while the remaining half were left on 50% reinforcement. It was hypothesized that the 50% treatment would prove more effective since the subjects would drop to 25% reinforcement if they persisted in the old habit, while the 100% subjects would continue to obtain the expected amount of reinforcement.

Since there was not a significant difference between treatments, the results were indecisive.

The present study is a renewal of the investigation. It differs from the other study in that the solution of the reversal problem requires a different concept from the one required in the original problem.

METHOD

Subjects: Thirty-two normal fifth grade public school children served in this study. The number of boys equaled the number of girls. These were divided into two groups consisting of eighteen and fourteen subjects respectively.

Apparatus: The equipment used consisted of a console having the subject's display panel on one side and the experimenter's control panel on the opposite side. The subject saw six indicator lights. A green light illuminated at the beginning of each trial. This served to advise the subject that a response was to be made. This light was in the upper left hand corner of the display panel. A second light (L_R) appeared in the upper right hand corner of the panel. This amber light automatically illuminated whenever a correct response was made in either the first or the second phase of the study. A third light (L_C) appeared on the top of the apparatus. This orange light came on only during the second and third stages of the experiment and then only when a reward (candy) was given.

Three red lights in vertical array (L_1, L_2, and L_3) were situated in the center of the panel. These were the discrimination lights and the subjects were required to observe these to determine which response should be made.

Below these lights was a three position toggle switch which could be moved laterally either left or right. It was spring loaded to the center position and consequently returned to that position whenever it was released. It was the task of the subject to observe the three discrimination lights and depending upon the combination illuminated, move the toggle switch either to the left (L) or to the right (R).

Eight possible combinations of three lights are possible. These are shown in Table 36.

Table 36

Example of One Problem given Showing the Correct Responses to each Variation of the Discrimination Lights

Stimulus Variation			Correct Response	
L_1	L_2	L_3	Stage 1 & Stage 2	Stage 3
on	on	on	L	R
on	on	off	L	L
on	off	on	R	R
off	on	on	L	L
on	off	off	R	L
off	on	off	L	R
off	off	on	R	L
off	off	off	R	R

The combination changed from trial to trial. This was accomplished by means of a 48 position multi-contact stepping switch which was driven by a solenoid. The circuitry was such that each combination appeared six times in each block of 48 trials. The order of presentation was random.

On the experimenter's panel, switches were provided to program the apparatus. In other words, the apparatus was constructed such that when a particular stimulus variation appeared, the correct response could be varied. When, for example, all lights were illuminated, the amber reinforcement light could be made to come on to either L or R.

The stepping switch was also under the experimenter's control. To start a trial, he would energize a push-button switch which advanced the program and set up the apparatus for the next trial.

Procedure: The experiment was conducted in three stages. In the first stage, 96 trials were given (the stepping switch went through two complete cycles). The program was always such that it could be solved on some rational basis. For example, the problem given for Stages 1 and 2 in Table 36 could be solved by the principle of, "Respond left when the center light is on and respond right when it is off."

During Stage 1, the amber light (L_R) came on each time a correct response was made. The subject was also given a piece of candy for each correct response. At the end of the first stage, the subject was told he would hereafter receive candy only when the orange light (L_C) came on, but that the amber light (L_R) would continue to illuminate each time a correct response was made. The apparatus operated such that L_C came on 50% of the time that L_R came on. In other words, in the second stage the subjects were dropped to 50% reinforcement. During this stage, the correct stimulus-response relations were identical with those obtained in Stage 1.

The third stage was ushered in by turning off L_R so that it would never illuminate. This was not brought to the attention of the subject. When he inquired (as he invaribly did) he was told that the light had burned out.

In the third stage, the problem solution was changed for all subjects. In the example shown in Table 36, the new solution might be described as "when L_1 and L_3 are the same, R is correct; but when they are different, L is correct." The difference was, of course, that the original responses were no longer appropriate to the new problem. For one group of subjects, L_C (the light associated with candy reinforcement) continued to illuminate on half of the correct responses. This group is called the 50% Reinforcement Group. For the second group of subjects, L_C was made to come on each time a correct response was made. In other words, they went back to 100% reinforcement on the new problem.

Both groups were again given 96 trials on this stage of the experiment and a record of errors was taken on each subject.

RESULTS

The outcome of Stage 3 was strongly in favor of the 100% Reinforcement Group. The mean errors for the 100% Reinforcement and the 50% Reinforcement Groups were 29.8 and 40.6 respectively. This difference is highly significant ($P < .01$) based upon the *t* test.

DISCUSSION

The hypothesis being tested was not substantiated. It is of interest to speculate as to the reason the superiority of the 100% Group was not evidenced in the third stage of the experiment described in Section 5. The basic difference between the studies was, of course, that in the earlier experiment a shift in stimulus relevancy was experienced but the basic concept (Similarity-Difference) remained. In the present study, a shift was made in the concept (this assumes, of course, that subjects solved the problems of the present experiment within a particular conceptual context) required for the solution.

The difference between the results of the two studies can be interpreted as evidence that in the solution of a problem of this kind, a dual process is involved. First, by induction, the subject seeks a general solution or concept. Second, by deduction and affirmation,

he validates or refutes his hypothesis. In the present study, the 50% Reinforcement Group was not given sufficient information in Stage 3 to induce a solution. In the earlier study, the original concept was still valid and it was only a matter of exchanging stimulus reference.

SUMMARY

The hypothesis was tested that subjects tend to seek new problem solutions when reinforcement falls below the expected amount. Subjects experienced a problem reversal under 100% reinforcement and under 50% reinforcement. The 100% reinforcement condition was superior. The hypothesis was not confirmed.

SECTION 7

THE EFFECT OF VARYING THE PROBABILITY AND THE AMOUNT OF REINFORCEMENT IN NORMAL AND RETARDED SUBJECTS

Conceived broadly, there are two ways in which a systematic reinforcement program can be administered. Between alternative responses, reinforcement can be made to vary in amount. Thus, the responses R_1, R_2, and R_3 can receive one, two, and three units of reinforcement respectively. Reinforcement can be money, tokens, or even score points as are given on an examination paper.

A second method is to vary the probability of reinforcement. By this method, the counterpart to the above system would be for R_1 to receive reinforcement 1/3 of the time, R_2 to receive reinforcement 2/3 of the time, and R_3 to receive reinforcement every time.

By either method, the relative values of R_1, R_2, and R_3 are equivalent since a subject can do three times as well by choosing R_3 as R_1 and twice as well by choosing R_2 in preference to R_1.

If the strength of a habit grows primarily as a function of the quantity and quality of reinforcement, the discrimination of R_1 from R_2 should proceed at about the same rate for both methods.

It is evident, however, that the amount of information given differs between the two methods. On the average, and over a large number of trials, there is three times as much information available on each trial by the amount method as is available by the probability method.

It is the intent of the present study to determine if the additional information made available by the amount method is of differential benefit to normal as compared to retarded subjects.

METHOD

Subjects: Ninety-one children served in this study. Of these, 64 were normal and twenty-seven were designated as retarded. Of the normal children, 32 were fourth and fifth graders (N_1) from the public schools while the remaining 32 were first and second graders (N_2). The retarded children (R) were obtained from both private and public sources. No attempt was made to control for diagnostic variations. Their ages ran from eight to sixteen. All were described as teachable or trainable. It was necessary in making up the group to eliminate several who had extreme difficulty in retaining interest in the test procedures. It was originally intended to include 32 retarded children in the experiment, but five of these subjects developed position habits (responded either to the left or right without regard to the stimulus presentation) and their scores are not included in the analysis.

<u>Apparatus</u>: The discrimination apparatus consisted of two hinged windows set into a panel facing the subject. The windows were spring loaded and could be opened by pushing against them. Behind the windows were small trays which held the reinforcement. Small pieces of cinnamon flavored candy (red hots) were used as reward. The windows were composed of two sheets of plexiglas set apart so that photographic transparencies could be inserted between them. A darkened background was provided behind the windows so that the pictures on the transparencies could not be seen until a light was illuminated. This light was turned off when the experimenter changed the pictures in the windows. A solenoid was used on each window to control the latch mechanism. The latches were engaged while the experimenter was changing stimuli and placing candy in the reinforcement trays. This prevented premature responses. The apparatus was contrived so that when the subject was required to respond, both windows were unlatched. When the subject opened one window, however, the latch on the opposite window automatically engaged preventing the subject from opening it as well. Subjects were sternly discouraged against opening both windows simultaneously.

The stimuli consisted of ten black and white photographic transparencies. Each of these were of a simple geometric figure. One contained a circle, one a triangle, another vertical lines, etc. On a given trial stimuli were placed in both windows. It was the subjects' task to open the window having either the greater amount of reinforcement or having the greater probability of reinforcement.

<u>Procedure</u>: Six subject-treatment groups were tested. These are shown in Table 37. For half of the groups (P) the probability of reinforcement varied depending upon the stimulus selected. For the other groups (A), the amount of reinforcement varied depending upon the stimulus. This is shown in Table 38.

Table 37

Subject Treatment Groups

Gp. N_1P 5th & 6th grades Probability Reinforcement N = 16	Gp. N_2P 1st & 2nd grades Probability Reinforcement N = 16	Gp. RP Retarded Probability Reinforcement N = 13
Gp. N_1A 5th & 6th grades Amount Reinforcement N = 16	Gp. N_2A 1st & 2nd grades Amount Reinforcement N = 16	Gp. RA Retarded Amount Reinforcement N = 14

Table 38

Reinforcement Schedule of the Probability and Amount Treatments for the Various Stimuli

	Probability Treatment		Amount Treatment	
Stimulus	Probability of Reinforcement	Amount when Reinforced	Probability of Reinforcement	Amount when Reinforced
A	90%	10 units	100%	9 units
B	80%	"	"	8 units
C	70%	"	"	7 units
D	60%	"	"	6 units
E	50%	"	"	5 units
F	40%	"	"	4 units
G	30%	"	"	3 units
H	20%	"	"	2 units
I	10%	"	"	1 unit
J	0%	"	"	0 units

The reinforcement schedules for stimuli are designated in a general way and do not refer to particular displays. The A stimulus might refer to the circle for one subject but to the triangle for another subject, etc.

It can be seen that in selecting Stimulus A, for example, a subject receiving the Probability Treatment had nine chances out of ten of receiving a reward. This is to say, that in the 40 times A appeared it was reinforced on 36 occasions. On each of these occasions, the reward consisted of 10 pieces of candy. On four of the times, however, no reward was placed behind Stimulus A. For a subject receiving the Amount Treatment, however, Stimulus A was always reinforced and the reward always consisted of nine pieces of candy. Similarly for the Probability Treatment, Stimulus G had ten units of reinforcement behind it 30% of the times it was presented, but had 3 units behind it each time it was presented under the Amount Treatment.

On each trial a stimulus was shown in each window. The subject selected by pushing against one of the windows. Two hundred trials were given each subject. Each stimulus was paired with each other stimulus four times and with itself twice. In this manner, each stimulus appeared twenty times during the first block of 100 trials and 20 times during the second block of 100 trials. The order of pair presentation was random within each block of trials.

RESULTS

Mean correct responses for groups and treatments are shown in Table 39. A correct response consisted of selecting that stimulus having either the greater probability of reward (as with the probability treatment) or the

greater amount of reinforcement (as with the Amount Treatment). Since out of the 200 trials there were only 180 responses which could be regarded as either correct or incorrect (twenty of the pairings were of the stimuli with themselves), the data is tabulated on the basis of 180 total discriminations for each subject.

Table 39

Mean Correct Responses for Groups and Treatments

Group N_1P M = 120.75	Group N_2P M = 100.25	Group RP M = 102.46	Probability Treatment M = 108.18
Group N_1A M = 136.50	Group N_2A M = 126.50	Group RA M = 91.86	Amount Treatment M = 119.43
Group N_1 M = 128.63	Group N_2 M = 113.38	Group R M = 96.96	Total M = 113.87

In general, the Amount Treatment (A) is superior to the Probability Treatment (P), although this is not the case for the Retarded Group whose mean, under the amount condition, is almost at chance level. The fourth and fifth graders (N_1) are superior to the first and second graders (N_2). These are in turn, superior to the retarded children (R).

These differences are evaluated in Table 40. Here it is seen that both primary effects (Reinforcement Procedure and Subject Source) are significant. The interaction of these effects is also significant.

Table 40

Analysis of Variance Attributable to Reinforcement (P and A) and to Subject Source (N_1, N_2, and R)

Source of Variation	SS	df	MS	F	P
A (Reinforcement)	2882.54	1	2882.54	12.10	<.01
B (Subject)	14692.46	2	7346.23	30.84	<.01
A x b	5372.47	2	2686.24	11.28	<.01
Within	20248.95	85	238.22		
Total	43196.42	90			

The significances of the differences between groups are provided by the *t* test analysis in Table 41.

Table 41

Results of *t* Tests on Group Differences

Group	N_1A	N_2P	N_2A	RP	RA
N_1P	2.86*	3.26*	0.92	2.91*	4.82*
N_1A		7.12*	1.98	7.17*	9.94*
N_2P			4.44*	0.38	1.51
N_2A				4.14*	6.26*
RP					2.02

* $P<.01$

DISCUSSION

Reinforcement plays two roles in behavior. It assuages a need or drive and it serves to direct subsequent responses. In its latter role, it confirms the appropriateness of the subject's action. Certain schedules of reinforcement supply more information than do other schedules, however. In the present instance, each reinforcement under the amount schedule logically contains ten times as much information as reinforcement under the probability schedule. This added information apparently was put to use by the normal subjects, but only tended to confuse the retarded subjects.

The most significant finding of this study was the interaction between Subject Source and Reinforcement Schedule. Normal subjects, even the very young, benefit when reinforcement is differentiated between stimuli with respect to amount. For retarded subjects, amount does not appear to play so significant a role. For these subjects reinforcement is reinforcement irrespective of amount.

Undoubtedly, many of the retarded subjects could not count as well as the normals could, but it probably is not merely a matter of counting. With the normals, the game was seen as one of adopting a strategy for maximizing their winnings. With the retarded subjects, it was largely a matter of getting candy, but they did not develop a criterion for success.

SUMMARY

Two reinforcement methods were compared using normal and retarded subjects. By one method, responses were differentiated in the amount of reinforcement received. By the second method, responses were differentiated by the probability of reinforcement.

It was found that the retarded subjects were inferior to the normals when compared by either reinforcement procedure. It was also evident that while the normals were facilitated by the amount of reinforcement procedure, the retarded subjects were adversely affected by it.

SECTION 8

EFFECT OF PARTIAL REINFORCEMENT ON NORMAL AND RETARDED SUBJECTS

The greater resistance to extinction resulting from partial reinforcement procedures as compared to 100% reinforcement has given rise to the opinion that the role of secondary reinforcement is one of information transmission (4). If secondary reinforcement is present, extinction is delayed (12). Some theorists (3) have adopted the notion that the stimuli present in a reinforcing situation become surrogates for primary reinforcement. As against this point of view, it can be proposed that the presence of these stimuli, which have been associated with primary reinforcement, serves instead to confirm the correctness of a response in the absence of primary reinforcement.

Regardless of which point of view is favored, the use of secondary reinforcement in the development of a habit implies that a mediational process is involved. Either the reinforcing properties of the stimuli must be built up through association with primary reinforcement, or the informational content of the stimuli must be established through this association. If the solution of a problem is dependent upon the development of this mediational process, then it seems reasonable to assume that it might affect learning in retarded subjects more than it would in normal subjects. The more tenuous the solution, the greater the difference between the performance of normal and retarded subjects should be.

METHOD

Subjects: Twenty-four normal (N) and forty-one retarded (R) children served as subjects. The age of the normal children was from eight to eleven years, while the age of the retarded subjects was from ten to eighteen. The normals were obtained from the third, fourth, and fifth grades of the public schools. The retarded subjects were obtained from the state institution for the mentally retarded. They were selected as some of the more "trainable" children available.

Procedure: The experiment was conducted by seating the subject across a table from the tester. On each trial, the experimenter showed the subject two stimuli. The subject was required to select one of the stimuli by pointing to it. The stimuli used consisted of five different figures. These are referred to as A, B, C, D, and E. Each of these stimuli were mounted on 5 x 5 inch cards. A was an equalateral triangle, B was a circle, C was a square, D was a cross, and E was a group of parallel lines.

In the one hundred trials given to a subject, each of the ten possible stimulus pairings was presented ten times. A was presented with B ten times, with C ten times, etc. The arrangement of pairs was balanced. For example, five of the B D presentations were with B on the left and five were with D on the left.

Correct responses were determined in the following manner: The response of pointing to A was always correct. B was always correct except when paired with A. C was correct except when paired with A or B. D was correct only when paired with E. The response of pointing to E was never correct.

The experimenter avoided ever telling the subject whether a response was correct or not. The subject could infer this from the reinforcement procedure, however. This procedure can be visualized from Table 42.

Table 42

Means and Size of Group for the Four Subject-Treatment Conditions

Reinforcement Procedure	Subjects: Normal	Subjects: Retarded
100% Primary Reinforcement and 100% Secondary Reinforcement	N = 12 M = 68.50	N = 23 M = 56.74
50% Primary Reinforcement and 100% Secondary Reinforcement	N = 12 M = 63.00	N = 18 M = 52.44

Twelve of the normal subjects received primary reinforcement (candy) 100% of the time for correct responses. On each correct trial, the subject was given a piece of candy. No candy was given if the choice was in error.

The other twelve normal subjects received only 50% primary reinforcement. In other words, the red poker chip was given on each correct trial and the blue chip on each incorrect trial, but candy was given on only half of the correct trials.

Twenty-three of the retarded subjects were given the 100% primary reinforcement condition and eighteen of these subjects were given the 50% reinforcement condition.

RESULTS

The mean correct responses for each group are shown in Table 42. As expected, the normal subjects were superior to the retarded subjects and the 100% reinforcement procedure was superior to the 50% reinforcement procedure. In Table 43, the significance of these differences is analyzed. The variance due to subject differences is significant ($P<.01$), but neither the reinforcement differences nor the interactions are significant.

Table 43

Analysis of Subject Source (Normal and Retarded) and Reinforcement (100% and 50%) Differences

Source of Variation	SS	df	MS	F	P
Reinforcement	272.18	1	272.18	2.76	NS
Subjects	1797.40	1	1797.40	18.26	<.01
S x R	95.56	1	95.56	.97	NS
Within	6005.88	61	98.46		
Total	8171.02	64			

DISCUSSION

It seems likely that with a larger N both subject and reinforcement differences would prove significant. In the absence of a larger interaction, however, an increase in the number of subjects does not seem to be justified.

There is no evidence on the basis of these data that controlled secondary reinforcement procedures affect normal and retarded subjects differentially.

SUMMARY

The hypothesis was tested that retarded subjects are more adversely affected by partial reinforcement procedures than are normal subjects. It was found that although normals were superior to retarded subjects on the discrimination problem given and that 100% reinforcement is more effective that 50% reinforcement, the interaction between subject category and reinforcement treatment was not significant. The hypothesis was not confirmed.

SECTION 9

CONCEPT FORMATION IN THE SOLUTION OF A THREE LIGHT PROBLEM

Concepts are unitary ways of organizing stimuli or the relationship between stimuli into response categories.

Any discrimination problem can be solved on a configurational basis. A relatively difficult problem, like the Similarity-Difference Problem, can be solved without the subject ever recognizing the similarity-difference principle. He can simply memorize the correct response for each combination of stimuli presented. There is a limit beyond which configurational learning cannot go, however. When the number of configurations become large, it is impossible for the subject to remember each correct response. He tends to search for a mnemonic device to aid him in his selection. In so doing, he may stumble upon a relationship that not only helps him remember the correct response for one stimulus configuration, but which is also applicable to another. This relationship is called a concept because it unifies two or more configurations. The subject no longer responds to a total configuration, but only that aspect which identifies it as belonging to the conceptual class.

In the present study, a number of problems were compared which differred in their potential for conceptual reference.

METHOD

Subjects: Twenty normal high school seniors served in this study. The group was composed of ten males and ten females. The school drew its students from a wide socio-economic range and the group can be regarded as a random sample from this population of high school students. It was the original intention to include grade school students as well as the mentally retarded in the experimental design. The results obtained with the high school subjects did not make this procedure advisable, however.

Apparatus: The equipment used consisted of a two panel console. The subjects panel was on the opposite side of the console from the experimenter's panel. There were five indicator lights and a switch on the subjects' side. The illumination of the starter light (L_S) indicated to the subject that a trial had begun and that he was to make a response. This light was located on the upper left hand corner of the panel. It remained illuminated until the subject had responded. A second light (L_R) was situated on the upper right hand corner. This light automatically illuminated whenever a correct response was made. Three red indicator lights (L_1, L_2, and L_3), situated in a vertical array in the center of the panel, served as the discriminanda. The illumination of these lights varied between trials. There are eight combinations of illumination and non-illumination possible with three lights. Each combination was used on each of the problems given.

Table 44

The Correct Stimulus-Response Relationships for the Twenty Problems

Discriminanda				Problem																			
C	L1	L2	L3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	R	R	L	R	L	R	L	R	L	R	L	R	L	L	R	L	R	R	L	R
2	0	0	X	R	R	L	R	L	R	L	R	L	R	R	L	R	R	L	R	L	L	R	L
3	0	X	0	R	R	L	R	R	L	R	L	R	L	L	R	R	R	L	R	L	L	R	R
4	0	X	X	R	R	R	L	L	R	L	L	R	L	R	L	L	L	R	L	R	R	L	L
5	X	0	0	R	L	L	L	R	L	R	R	L	L	L	R	L	L	L	R	L	L	R	L
6	X	0	X	R	L	R	L	L	L	L	R	R	L	L	L	R	L	R	L	L	L	L	R
7	X	X	0	R	L	R	L	R	R	R	L	R	R	R	R	L	R	R	R	R	L	L	L
8	X	X	X	L	L	R	R	R	L	R	L	L	R	R	L	L	R	L	L	R	L	R	R

C - light combinations

0 - light off
X - light on

L - left response correct
R - right response correct

Table 45

Comparison of Errors by Duncans Multiple Range Test

P*	Means	1	3	16	4	19	7	17	15	20	9	8
	Means	5.2	7.0	10.4	10.5	10.8	11.1	11.7	12.5	12.9	13.1	13.4
1	5.2		1.8	5.3	5.4	5.7	5.9	6.5	7.4	7.8	8.0	8.2
3	7.0			3.5	3.6	3.9	4.1	4.7	5.6	6.0	6.2	6.4
16	10.4				.1	.4	.7	1.3	2.1	2.5	2.7	3.0
4	10.5					.3	.6	1.2	2.0	2.4	2.6	2.9
19	10.8						.3	.9	1.7	2.1	2.3	2.6
7	11.1							.6	1.5	1.9	2.1	2.3
17	11.7								.9	1.3	1.5	1.7
15	12.5									.4	.6	.9
20	12.9										.2	.5
9	13.1											.3
8	13.4											
6	14.0											
5	14.5											
18	14.7											
2	14.8											
13	14.9											
11	15.0											
12	15.1											
14	17.3											

P*	6	5	18	2	13	11	12	14	10	Shortest Significant Ranges
	14.0	14.5	14.7	14.8	14.9	15.0	15.1	17.3	17.7	
1	8.8	9.4	9.5	9.6	9.8	9.9	10.0	12.2	12.6	$R_{2} = 4.9$
3	7.0	7.6	7.7	7.8	8.0	8.1	8.2	10.4	10.8	$R_{3} = 5.1$
16	3.6	4.1	4.3	4.4	4.5	4.6	4.7	6.9	7.3	$R_{4} = 5.3$
4	3.5	4.0	4.2	4.3	4.4	4.5	4.6	6.8	7.2	$R_{5} = 5.4$
19	3.2	3.7	3.9	4.0	4.1	4.2	4.3	6.5	6.9	$R_{6} = 5.5$
7	2.9	3.5	3.6	3.7	3.9	4.0	4.1	6.3	6.7	$R_{7} = 5.6$
17	2.3	2.9	3.0	3.1	3.3	3.4	3.5	5.7	6.1	$R_{8} = 5.7$
15	1.5	2.0	2.2	2.3	2.4	2.5	2.6	4.8	5.2	$R_{9} = 5.8$
20	1.1	1.6	1.8	1.9	2.0	2.1	2.2	4.4	4.8	$R_{10} = 5.8$
9	.9	1.4	1.6	1.7	1.8	1.9	2.0	4.2	4.6	$R_{11} = 5.8$
8	.6	1.2	1.3	1.4	1.6	1.7	1.8	4.0	4.4	$R_{12} = 5.9$
6		.6	.7	.8	1.0	1.1	1.2	3.4	3.8	$R_{13} = 5.9$
5			.2	.3	.4	.5	.6	2.8	3.2	$R_{14} = 6.0$
18				.1	.3	.4	.5	2.7	3.1	$R_{15} = 6.0$
2					.2	.3	.4	2.6	3.0	$R_{16} = 6.0$
13						.1	.2	2.4	2.8	$R_{17} = 6.0$
11							.1	2.3	2.7	$R_{18} = 6.1$
12								2.2	2.6	$R_{19} = 6.1$
14									.4	$R_{20} = 6.1$

* Problem

It was the task of the subject, on a trial, to observe the light combination and decide on the proper response.

Immediately below the discriminanda a three position toggle switch was located. It was spring loaded to the center position. Subjects registered their responses by moving the switch either to the left (L) or to the right (R). If the response was correct, L_R would immediately illuminate.

The correct responses with reference to the light combinations could be varied by means of eight switches on the experimenter's control panel.

The order of light combination presentations was fixed. A 48 position multi-contact stepping switch was used to set up this order. After each trial, the switch would advance one position setting up a different combination of lights. On the 48 positions of the stepping switch, each of the eight possible light combinations occurred six times. The order of presentations was random.

Procedure: A 20 x 20 latin square design was followed. This is to say that twenty subjects each received the same twenty problems. The order of problem presentation differred between subjects such that a particular problem was presented to only one subject at a given stage in the overall order of presentation.

The problems are shown in Table 44. Problem number 1, for example, could be solved by making an R response to each combination of lights presented except that combination in which all of the lights were illuminated. Problem number 3 could be solved by an R response whenever two or more lights were illuminated, etc.

Except for problems 1, 13, and 18, the problems were a random sample from all possible problems in which there are four left responses and four right responses.

Assuming that subjects use concepts in the solution of problems of this kind, then some of the problems should be more easily solved than others. Number 20, for example, can be solved by the concept, "If L_1 and L_3 are identical, R is correct; but if they are different, L is correct." Number 12 can be solved by, "If L_3 is illuminated, L is correct; but if it is not illuminated, R is correct." Other problems are much more difficult to conceptualize upon. For some, in fact, it is probably necessary to memorize each combination-response relation separately without the aid of conceptual organization.

A total of 96 trials (12 of each combination) were given on each problem. A record was kept of the number of errors made by each subject on each problem.

RESULTS

The mean errors on each of the problems are presented in the left

hand column of Table 45. These are listed in order from the easiest to the most difficult. Problem number 1 was easiest, number 3 next, etc. The differences between the problems are analyzed in Table 46 along with the differences between subjects and the differences between orders of presentation. All of these differences are significant at the 99% level of confidence. The significances of the differences between specific problems are presented in Table 45. In general, problems 1 and 3 were significantly easier than the other problems while 14 and 10 were significantly more difficult. The differences between most of the problems are not significant.

Table 46

Analysis of Variance Contributed by
Problem, Subject, and Order of Presentation Differences

Source of Variation	SS	df	MS	F	P
Problems	3659	19	192.58	3.11	$<.01$
Subjects	7980	19	420.00	6.78	$<.01$
Order	11718	19	616.74	9.96	$<.01$
Residual	21185	342	61.94		
Total	44542	399			

DISCUSSION

There is evidence that the subjects did conceptualize upon the problems. Problems 1 and 3 which were significantly superior to most of the other problems are relatively easy to organize conceptually. The most difficult problems (14 and 10) are particularly difficult to organize as concepts. In the aggregate, however, there was not a very good relationship between errors and the potential for conceptual organization. Subjects apparently engage in a great deal of limited conceptualization. They may break a problem into two or more less general concepts that could be solved in terms of one concept.

Although it was originally the plan to test retarded subjects on these problems, this approach was abandoned. Since there was little difference between most of the problems using normal subjects, there was little hope that an interaction between intelligence and problems could be found.

SUMMARY

Twenty normal subjects were each required to solve twenty problems dealing with the variation in three lights. While it was found that some problems were more easily solved than others, there was little evidence that subjects tend to organize material of this kind into unitary concepts.

SECTION 10

COMPARISON OF THREE METHODS OF REINFORCEMENT QUANTIFICATION

The present study relates to the problem of stimulus synthesis. The perceptual process by which different cues are combined or integrated is difficult to ascertain experimentally. It is evident, however, that in almost any perceiving situation a multiplicity of stimulus cues are afforded. On the basis of these numerous cues the human often, unconsciously or without noticeable effort, obtains a unified impression which we call a "percept" or a "concept". In the present context no distinction is made between these terms. The act of perceiving a cow or of having a direct experience which is automatically and without reservation directly related to the concept of "cowness", are viewed as different descriptions of the same process.

In the rarified atmosphere of the experimental laboratory, subjects often have occasion to make such judgments as "if black, go left; but if white, go right." Most real life discriminations are not as simple as this. To get candy from a refractory parent is a more complex task than manipulating the usual discrimination apparatus, or Skinner Box. However, children develop very satisfactory records in tasks of this kind. In examining the situation, we can realize that the probability of obtaining candy upon request varies with the parent approached, the time of day, the manner in which the request is tendered, and a host of other factors as well as their more subtle interactions. The child does not, of course, analyze these probabilities mathematically, but nevertheless, he does develop, by means of some perceptual process, a type of behavior which approximates a strategy based upon the objective probability values.

Other discriminations may be more simple. A child obtains an experience and automatically categorizes that experience as a cow. Cow experiences are, however, closely related to horse experiences. There is considerable overlap in the variations exhibited by the two species. Size, weight and color may be useful cues, but are never final. Horn and cloven hoof may provide absolute criteria, but discrimination can still be made when these are invisible. Undoubtedly, a large number of cues go into such a concept as "cowness" or "horseness".

It is the intent of this experiment to study how different cues are put together into a unitary percept. Since it was impossible to control for the past experience of the subjects, it was decided to create two artificial populations of things rather than use familiar categories. A population of X's and Y's were assembled. Rather than having a multitude of variations, as is often encountered in real life, two dimensions of experience were varied. This was deemed necessary in order to simplify analysis. There was an overlap between the two categories in terms of the variations shown. In certain extreme

instances the stimulus condition used was invariantly the property of one or other of the two categories. More often, however, a stimulus variation gave something less than unity, and more than zero probability of belonging to a particular category.

The method of determining the perceptual process was indirect. The procedure was based upon the following notion: Since reinforcement gives the subject information concerning the appropriateness of his response, a reinforcement system which varies quantitatively and proportionately to the perceptual process will then prove more effective than a reinforcement system which is not proportional to the perceptual process. Taking an analogy from linear frequency theory, it was assumed that the subject and the reinforcement schedule could be represented as components within a closed loop system, as shown in Figure 2. The subject acts as an integrator (combines the cues in a certain way) of the output from the stimuli. The integrative function is subject to modification from the feedback provided by the reinforcement. If the subject's output is proportional to the feedback signal, the system then becomes stable in the sense that the integrative function no longer undergoes change. On the other hand, if they are not proportionate to one another, the system undergoes mutation in "search" of stability.

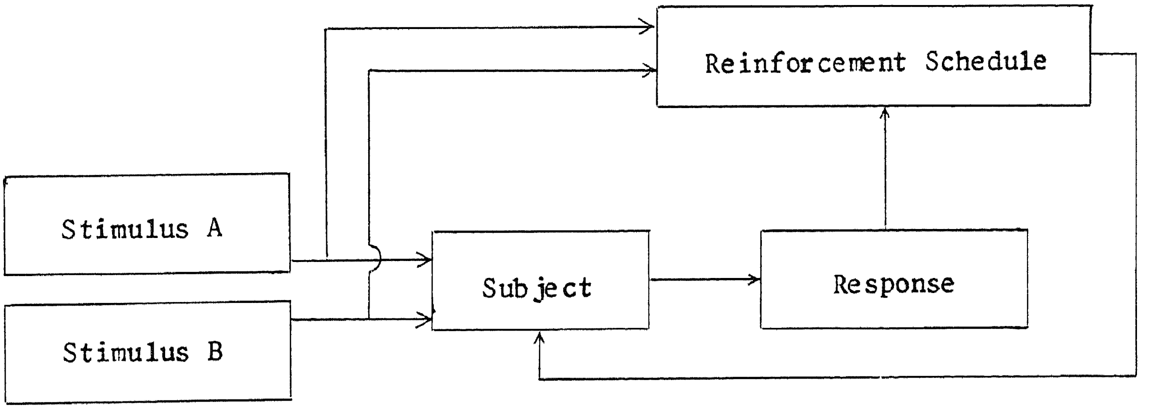


FIGURE 2. Diagramatic Representation of the Problem.

As shown, the amount of reinforcement is determined by the reinforcement schedule on the basis of the stimuli presented and of the response elicited. If the reinforcement schedule is based upon an integrative function identical with that which the subject is using, the system can become stable. If, however, the functions differ, stability cannot be attained. It is assumed that with instability, errors increase.

Since the integrative function of the subject is unknown, an approximation may be obtained by manipulating the reinforcement schedule until stability is achieved.

In the present experiment three integrative functions are examined. It is not proposed that any one of the functions must necessarily be the one used by the subjects. They were, however, selected as being broadly representative of the possibilities for stimulus integration.

METHOD

Subjects:

One hundred and sixteen normal public school children were tested. Seventy-two subjects were from sixth grade classes, and 44 subjects were from high school classes.

There were three treatments (reinforcement schedules), and consequently six subject-treatment groups. The number of subjects per group is shown in Table 47.

Table 47.

Subject Treatment Groupings

	High School	Grade School
Linear Hypothesis	N = 10	N = 24
Multiplicative Hypothesis	N = 14	N = 23
Square Root Hypothesis	N = 20	N = 25

Apparatus:

The equipment used consisted of a 35mm slide projector, a projector screen, and 420 color transparencies. All of the pictures, which constituted the stimuli in the experiment, consisted of two objects photographed simultaneously. One of the objects (A) was a large glass graduate similar to the type used by chemists. In the photographs it contained a milky white fluid. The amount of fluid in the glass varied between pictures. In the 420 pictures there were 21 variations in the fluid quantity. Twenty pictures were obtained with the graduate containing no fluid, twenty with one unit of fluid, twenty with two units, etc. The final twenty pictures showed twenty units of fluid.

Also present in each photograph was a marble board (B). It was a blue tray having a 5 x 5 matrix of holes into which red marbles could be placed.

The number of marbles present varied between pictures. The number of variations again amounted to twenty-one. For instance, twenty of the pictures showed no marbles, twenty showed one marble, twenty showed two marbles, etc.

In photographing these stimuli and their variations, the association of fluid quantity (A) with marble quantity (B) is carefully controlled. Two populations were created from the 420 pictures. Of these, 210 were called X's and the remaining 210 called Y's. The X and Y populations differed in their respective distribution of the A and B variations. For example, X contained a greater number of pictures showing the graduate full than did Y. On the other hand, Y contained more pictures showing a larger number of marbles than did X. The distribution of X's as a function of variation A is shown in Figure 3. It will be seen that as A increases, the frequency of X's also increase. In Figure 4 the distribution of Y's as a function of the variation in A is shown. Since the distribution of X's and Y's taken together is rectilinear with respect to A, the Y distribution compliments the X distribution.

In Figure 5 the distribution of X's is shown as a function of the variation in B. Thus, as the number of marbles increases, the frequency of X's decreases. Again, since the distribution of X's and Y's taken together is rectilinear, the distribution of Y's as a function of the variation in B complements the X distribution.

The association of the A and B variations in the picture was random except for the restrictions imposed by the distributions. For example, in determining the actual pictures that were to make up the X population, one slip of paper bearing the number "1", two slips each bearing the number "2", etc., were labeled and placed in a container designated A. Twenty slips of paper bearing the number "0", nineteen slips of paper bearing the number "1", etc., were then made, and all of these were consigned to a second container labeled B.

A slip of paper randomly drawn from one container was then paired with a slip of paper similarly drawn from the other container. Pairing was continued in this manner without the replacing of the slips of paper until the 210 slips were exhausted from each container. The photographs composing the X population were based upon the stimulus variation pairings obtained.

The same procedure was enacted in the generation of the population of Y pictures, except that the distributions differed as explained above.

The correlation of A to B for both X and Y was $r = -.39$.

The 420 pictures were then assembled in the order of presentation which was to be given the subjects. This order was randomized except that in each block of 42 pictures there was an equal number of X's and Y's. Each variation of A and B was also represented twice.

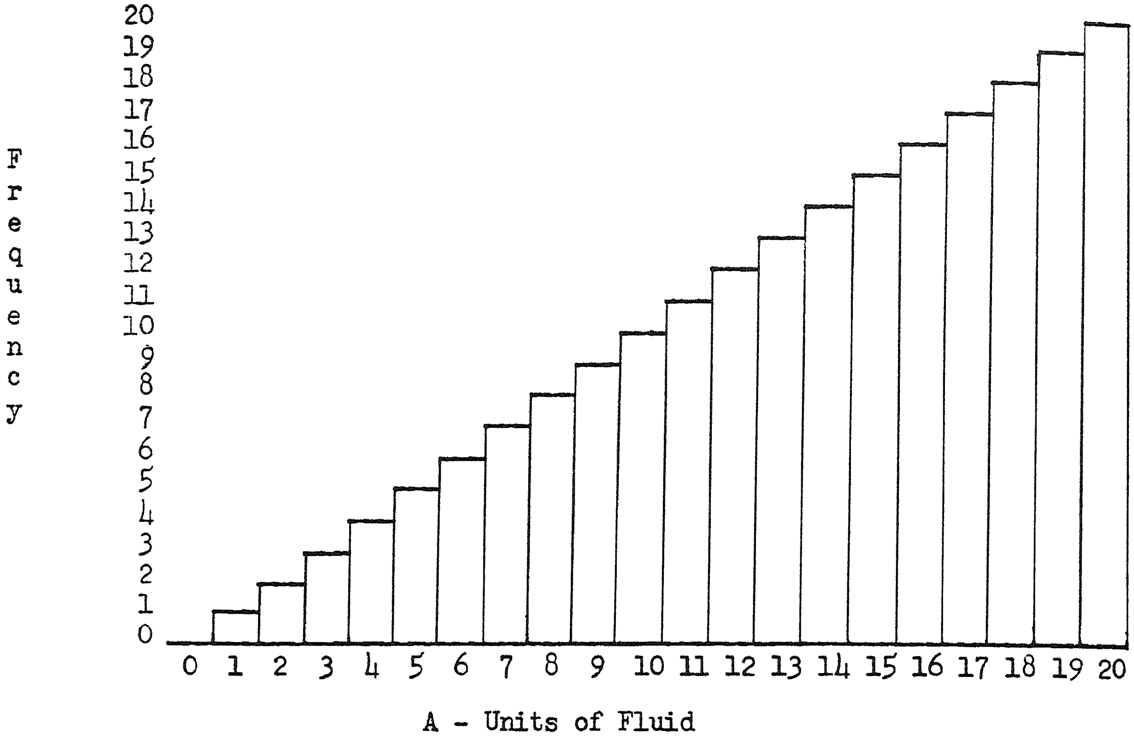


Figure 3

Frequency of X as a Function of the Variation in A

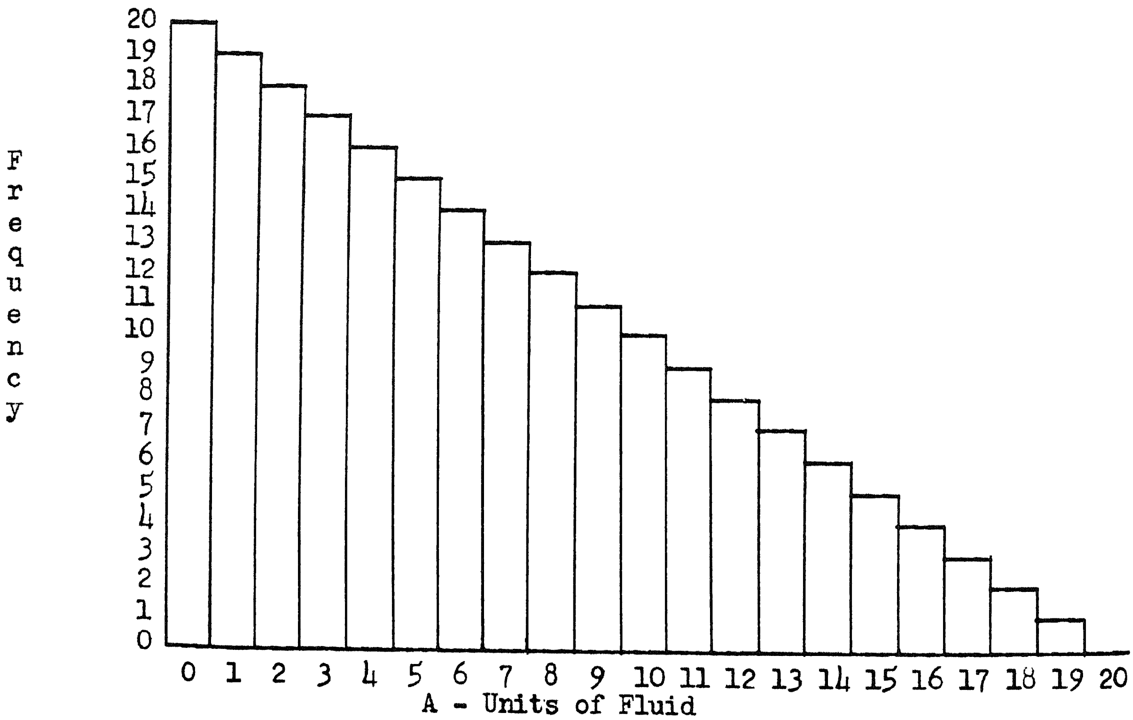


Figure 4

Frequency of Y as a Function of the Variation in A

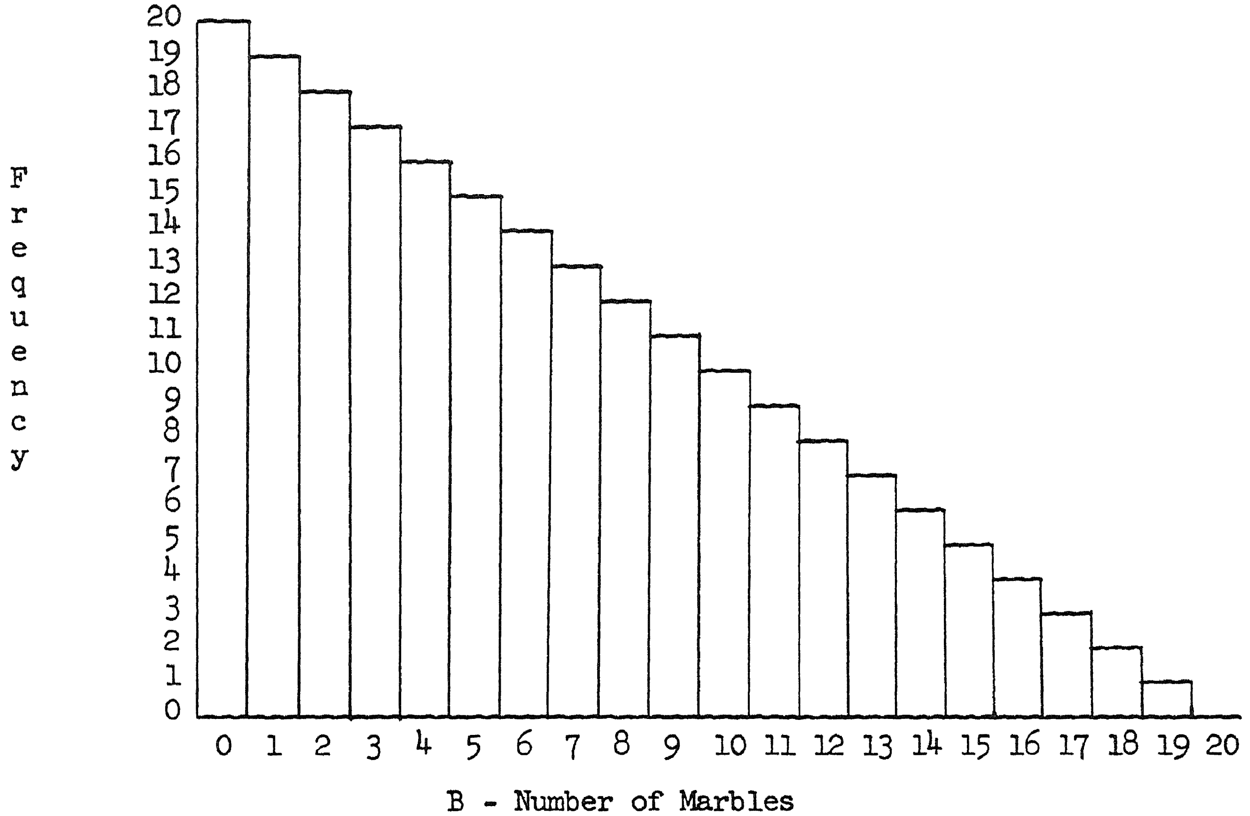


Figure 5

Frequency of X as a Function of the Variation in B

Procedure :

Six subject-treatment groups were tested. Each group was tested as a unit and received approximately an hour of testing on each of four consecutive days. The procedure followed was to project a picture on the screen. The subjects were allowed five seconds to look at the picture before it disappeared. They were then required to write in their answer sheets whether they thought it was an X or Y. After all of the answers were recorded, each subject passed his pencil to a designated neighbor, who was also a subject. The arrangement for pencil passing was such that each pencil went through each subject's hand before returning to its point of origin. It was also arranged so that the color of the ink in the pencils would alternate from red to blue between stimulus presentations. Thus, on the answer sheet an answer written in red would be followed by an answer written in blue.

After the pencils had been passed and the testers were confident that all subjects had recorded their answer, the subjects were given information (reinforcement) relating to the correctness of their response. For example, they might be told, "If you marked the answer an X, score yourself 725 points, but if you marked it a Y, score yourself 275 points." The pencil passing procedure was a strategum designed to prevent a subject from delaying his response until after the information was given.

The subjects were never told whether the picture they had just seen belonged to the X or Y stimulus populations. They were only told the value of their answers. If an X response was worth more than a Y response, the picture was probably an X. This was not necessarily the case, however. An understanding of this statement may be facilitated if the method by which the values were obtained is examined.

The set of answers differed between treatments. One treatment (R1) was based upon the assumption that the process of perceptual integration is essentially additive. In other words, when a concept (such as the concept of "Xness") is founded upon two distinguishing cues (A and B), the synthesis can be approximated by the formula,

$$P_X = \frac{P_A + P_B}{P_A + P_B + (1-P_A) + (1-P_B)}$$

Suppose, for example, a picture is shown in which there are five units of fluid and nine marbles. Referring to Figures 3 and 5, it can be seen that the probability of the picture being an X on the basis of A is 5/20 = .25, but that the probability of it being an X on the basis of B is 11/20 = .55. The above formula simply states that the probability of the picture being seen as an X (P_X) based upon the integration of A and B is the sum of their separate probabilities (P_A and P_B) divided by this sum, plus the probability of it being seen as a Y on the basis of A ($1-P_A$), plus the probability of it being a Y on

the basis of B $(1-P_B)$. In the present example it is hypothesized that the picture will be seen as a Y since:

$$P_X = \frac{.25 + .55}{.25 + .55 + .75 + .45} = .40$$

$$P_Y = 1.00 - .40 = .60$$

It is reasonable to assume that since one of the functions of reinforcement is to give information concerning the correctness or incorrectness of a response, a reinforcement system which is quantitatively proportional to the synthesizing process will provide the most efficient feedback system for correction. If the synthesizing process can be represented by the above formula, the most efficient reinforcement procedure (that which would give the best information) should follow the same formula. Let us say that the strength of a percept may vary from zero to 1000, where zero means that the subject is absolutely certain that it is not an X, where 1000 means that he is absolutely certain that it is an X, and where 500 means complete uncertainty. If the subject has a percept whose value is 800 and receives reinforcement (which is also based upon a scale running from zero to 1000) of an equivalent amount, then the percept remains stable. On subsequent presentations of the same stimulus, the strength of the percept will again be 800. On the other hand, if the strength of the percept is 800 but the reinforcement is 900, the percept will undergo modification. Perceptual readjustment will be made in the direction indicated by the amount of reinforcement obtained.

A perceptual-reinforcement system in which the perceptual process differs from the reinforcement process (in the sense that they can never come into full agreement) is inherently unstable and is more subject to error than when the processes are identical. Following this argument further, one could say that the perceptual process may be identified by varying the reinforcement process until the minimum error condition is obtained.

Admittedly, there are difficulties in testing a hypothesis of this kind. Even if the synthesizing process remains constant through time, it is evident that the response strengths, based upon the individual elements, do not remain constant. At the beginning of testing the subject has no way of differentiating between X and Y on the basis of any particular variation in A. The mean value of P_A would be equal to .50 on the first trial. As testing progresses, however, P_A changes toward an approximation of the true probability. Since the momentary value of P_A is unknown, the synthesizing process cannot be identified until a certain degree of stability emerges.

A second reinforcement procedure (R2) was calculated on the basis of the formula,

$$P_X = \frac{P_A \ P_B}{P_A P_B + (1-P_A)(1-P_B)}$$

This is designated as the multiplicative hypothesis. It is based on the assumption that the integrative process is one in which the extreme values of P_A and P_B are magnified at the expense of the more central values. If the subjects tend to focus their attention primarily on those stimulus aspects which deviate (either physically or probabilistically) greatly from the normal, then this formula is a better predictor of the strength of the percept than the linear hypothesis would be. It is also true that a reinforcement schedule based upon the multiplicative hypothesis would be superior.

In the example given earlier, P_X would obtain a smaller value than was shown earlier:

$$P_X = \frac{.25 \times .55}{(.25 \times .55) + (.75 \times .45)} = .29$$

P_Y would be greater than before:

$$P_Y = 1.00 - .29 = .71$$

The final reinforcement treatment (R3) was based upon the assumption that subjects tend to overemphasize the value of those stimuli having medium probability values at the expense of those having either very high or very low probability values. It was based upon the following formula:

$$P_X = \frac{\sqrt{P_A + P_B}}{\sqrt{P_A + P_B} + \sqrt{(1-P_A + (1-P_B)}}$$

This procedure generally results in a value nearer .50 than is the case with either of the other formulas:

$$P_X = \frac{\sqrt{.25 + .55}}{\sqrt{.25 + .55} + \sqrt{.75 + .45}} = .45$$

$$P_Y = 1.00 - .45 = .55$$

The P_X and P_Y values for each of the 420 pictures, using each of the hypotheses, were calculated before testing. These values were read to the subjects after each response. The values calculated for the linear hypothesis were read to the R1 subjects, the values calculated for the multiplicative hypothesis were read to the R2 subjects, and the values calculated for the square root hypothesis were read to the R3 subjects.

RESULTS

In order to observe if changes in stimulus synthesis occur as testing progresses, the data have been broken down into 10 blocks of 42 trials each. In Table 48 the mean correct responses are presented for each block of 42 trials. A correct response was defined as the response which yielded the largest number of score points. If, for a given group and picture, P_X was greater than P_Y, then X is regarded as the correct response.

A summary of the analysis of the difference between treatments and subject sources is presented in Table 49.

In the first block of 42 trials, the difference between high school and grade school subjects is not significant, although the difference favors the high school subjects. The differences between reinforcement schedules are significant, with the multiplicative hypothesis being most favored and the square root solution least favored. The interaction between hypotheses is significant and is presumably due to the high scores made by the high school subjects on the multiplicative solution (R2).

In the second block of trials both primary effects are significant. This effect is due to the superiority of the high school students, and also to the R2 treatment.

In the third block of trials only the interaction is significant. This finding suggests that while R2 remains the best solution for the high school subjects, the grade school subjects having R3 are improving rapidly.

For the remaining blocks of trials an interaction is always evident, although it does not always achieve statistical significance. This interaction is generated out of the relative superiority of the grade school subjects on R3 and the inferiority of the high school subjects on the same solution.

One final trend merits discussion: the initial results favored R2 for both subject sources. The trend for the high school subjects is toward the R1 treatment, although its superiority over R2 is never demonstrated. Superiority for the grade school subjects is from R2 to R1, and finally to R3.

DISCUSSION

Learning was facilitated for both subject sources by the R2 treatment in the early stages of testing. This is interpreted as meaning that early in testing subjects tend to overemphasize the very high and very low probability values. This would, of course, favor the multiplicative hypothesis. As training progresses, however, greater emphasis is placed upon the more central probability values. Thus the grade school subjects displayed superior performance for the R3 treatment. The relative advantage of the R2 treatment for the high school subjects also diminished in time, and finally the R1 treatment seemed to prove most favorable for this group.

Table 48.

Mean Correct Responses for Each Block of 42 Trials

TRIALS	RP	G	H
1 - 42	R1	25.00	24.00
	R2	25.65	32.64
	R3	21.76	22.60
43 - 84	R1	29.67	31.00
	R2	30.00	35.50
	R3	22.76	27.30
85 - 126	R1	32.63	35.20
	R2	31.09	35.93
	R3	31.80	29.75
127 - 168	R1	35.46	38.60
	R2	32.48	37.21
	R3	33.84	31.70
169 - 210	R1	35.75	39.10
	R2	34.57	38.36
	R3	35.84	34.05
211 - 252	R1	35.67	38.50
	R2	34.87	37.36
	R3	36.52	33.80
253 - 294	R1	36.00	39.60
	R2	35.17	38.14
	R3	37.00	34.90
295 - 336	R1	37.63	40.20
	R2	35.70	39.93
	R3	37.68	35.90
337 - 378	R1	36.50	37.50
	R2	34.78	37.57
	R3	36.88	35.30
379 - 420	R1	38.21	38.90
	R2	37.00	38.43
	R3	38.52	36.50

G -- Grade School Subjects
H -- High School Subjects
RP - Reinforcement Program
R1 - Linear Reinforcement Program
R2 - Multiplicative Reinforcement Program
R3 - Square Root Reinforcement Program

Table 49.

Summary of the Analyses of Variance

TRIALS	*	SS	df	MS	F	P	TRIALS	*	SS	df	MS	F	P
1 - 42	S	112.58	1	112.58	3.58		211-252	S	2.33	1	2.33	<1	
	R	772.52	2	386.26	12.29	<.01		R	27.38	2	13.69	<1	
	SxR	327.62	2	163.81	5.21	<.01		SxR	190.39	2	95.20	2.79	
	W	3457.79	110	31.43				W	3753.10	110	34.12		
	T	4670.51	115					T	3973.20	115			
43-84	S	311.09	1	311.09	6.82	<.05	253-294	S	22.95	1	22.95	<1	
	R	1177.79	2	588.90	12.91	<.01		R	20.04	2	10.02	<1	
	SxR	193.74	2	96.87	2.12			SxR	194.24	2	97.12	3.11	<.05
	W	5017.59	110	45.61				W	3439.22	110	31.27		
	T	6700.21	115					T	3676.45	115			
85-126	S	33.49	1	33.49	<1		295-336	S	34.95	1	34.95	1.23	
	R	143.49	2	71.75	1.71			R	44.58	2	22.29	<1	
	SxR	264.01	2	132.01	3.14	<.05		SxR	202.99	2	101.50	3.57	<.05
	W	4621.73	110	42.02				W	3124.26	110	28.40		
	T	5062.72	115					T	3406.78	115			
127-168	S	31.75	1	31.75	<1		337-378	S	5.27	1	5.27	<1	
	R	236.74	2	118.37	3.13	<.05		R	16.59	2	8.30	<1	
	SxR	284.00	2	142.00	3.75	<.05		SxR	97.21	2	48.61	2.33	
	W	4164.02	110	37.85				W	2298.68	110	20.90		
	T	4716.51	115					T	2417.75	115			
169-210	S	37.09	1	37.09	<1		379-420	S	2.01	1	2.01	<1	
	R	56.68	2	28.34	<1			R	16.55	2	8.28	<1	
	SxR	202.86	2	101.43	2.10			SxR	64.46	2	32.23	1.45	
	W	5316.58	110	48.33				W	2445.53	110	22.23		
	T	5613.21	115					T	2528.55	115			

*S - Subject Source
R - Reinforcement Schedule

*W - Within Variance
T - Total Variance

The most interesting feature in these data is the interaction between subject source and reinforcement treatment. For some reason the grade school subjects were superior to the high school subjects with the R3 treatment, although they were inferior on the other treatments. One would normally expect the older subjects to make the finer discriminations involved in the R3 treatment.

SUMMARY

Subjects drawn from two sources (high school and grade school) were required to discriminate between two artificial populations of stimuli which varied along two dimensions. Three reinforcement treatments were used: one was best adapted to the subject who gave proportionate emphasis to the high and low values of probability occurrence (R1), another reinforcement system was most useful to the subject who tended to overemphasize the high and low probability values (R2), and a third reinforcement system favored the subject who gave disproportionate emphasis to the low probability values (R3).

In the aggregate, R2 proved most useful early in testing. However, as training progressed, the R3 treatment was most useful for the grade school subjects.

SECTION 11

PERCEPTUAL SYNTHESIS INVOLVING TWO STIMULUS VARIATIONS

The rational for this study is similar to that described in the preceeding section. In common with the other study its purpose was to identify or to describe the synthesizing process by which subjects integrate stimulus cues in the formation of a percept. In the present experiment, however, reinforcement was not varied quantitatively. Subjects were simply told if they were right or wrong in their selection.

The procedure was based upon the notion that the strength of the percept arising from a stimulus situation will vary between subjects and that the number of subjects who perceive the stimulus as belonging to a particular conceptual class will be proportional to the mean strength of the percept. It is also assumed that if the mathematical equivalent to the perceptual process is known, it should correlate highly with the number of subjects perceiving the stimulus as belonging to a given class of objects. Of two mathematical expressions the one which more nearly approximates the perceptual process will correlate most highly with this proportion of subjects.

METHOD

Subjects: In all, 115 subjects were used in this experiment. All were normal and were obtained from the public schools. Of these, 56 were 5th grade students and the balance were from high school.

Procedure: By means of a slide projector, the subjects were shown a series of 420 Kodachrome transparencies. The pictures were those used in the previous study. In each photograph, there was a glass graduate (which varied in the quantity of fluid it contained) and a marble board (which also varied in the number of marbles shown). Two hundred and ten of these pictures were called X's and the remaining were called Y's. The distribution of X's as a function of the variation in A (the quantity of fluid in the graduate) is shown in Figure 3, while the same distribution as a function of the variation in B (number of marbles) is shown in Figure 5.

Three treatment groups can be delineated. Stimulus A Group, which consisted of 15 grade school subjects, saw only the glass graduate. This was accomplished by masking out the marble board so that it could not be seen. Stimulus B Group, which consisted of 18 grade school subjects, saw only the marble board. The Two Stimulus Group, which consisted of 59 high school and 23 grade school subjects, were allowed to witness both the graduate and the marble board.

Each group received the same order of picture presentation. A trial consisted of being shown one of the pictures for 5 seconds. A subject was then required to write in his answer sheet whether he judged

the picture to be an X or a Y. Pencils were then exchanged (the manner in which this was accomplished is explained in the previous section) between subjects in order to prevent postponed answers being recorded. The correct answer was then given. Unlike the previous procedure (where quantitative values were placed upon answers), the subjects were told only whether the picture was an X or a Y.

It should be understood that, although the pictures were the same as those used previously, the answers given did not always agree with what could be inferred from the answers given in the previous study. In the foregoing study, subjects were not told that a picture was actually an X or a Y. Rather they were told the value of their judgements as determined on the basis of certain assumptions concerning the perceptual process. When consideration is given to the distribution of X and Y on the basis of the A and B variations, it can be appreciated that there are numerous examples in the X population which have more the appearance of Y's than X's. It is also true that many Y's are more properly judged as X's.

RESULTS

The subjects receiving stimulus A alone were tested as a group, as were the subjects receiving stimulus B alone. The subjects receiving both stimuli, however, were made up of five test groups. In order to determine whether there was any difference of significance between the two stimulus groups, an analysis of variance was conducted on each block of 42 trials. None of the variances attributable to grouping were significant and the results are not reproduced in this report. Since the differences were without significance, the groups receiving the Two Stimulus Problem were treated as a single group in all subsequent analysis.

Six different hypotheses concerning the act of perceptual integration were analyzed. These are described below.

1. $P_X = P_A$

This hypothesis states that the probability that a two stimulus presentation would be perceived as an X is equivalent to the probability that it would be seen as an X on the basis of the A variation alone. In other words, it states that the B variation is ignored. It is referred to as the A hypothesis (H_A).

2. $P_X = P_B$

This hypothesis states that the A variation will be ignored. It is referred to as the B hypothesis (H_B).

3. $$P_X = \frac{P_A + P_B}{P_A + P_B + (1-P_A) + (1-P_B)}$$

This hypothesis states that the act of synthesis is additive or proportionate to the probabilities based upon the separate variations. It is referred to as the linear hypothesis (H_{A+B}).

4.
$$P_X = \frac{P_A\, P_B}{P_A\, P_B + (1-P_A)(1-P_B)}$$

This hypothesis states that in perceptual integration the high and low probability values are given disproportionate emphasis. It is referred to as the multiplicative hypothesis (H_{AB}).

5.
$$P_X = \frac{\sqrt{P_A + P_B}}{\sqrt{P_A + P_B} + \sqrt{(1-P_A) + (1-P_B)}}$$

This hypothesis states that the mid-range probability values are magnified. It is called the square root hypothesis ($H_{\sqrt{A+B}}$).

6. $P_X = P_A$ or P_B whichever deviates the farthest from $P = .50$

This hypothesis states that subjects base their judgement entirely on either one or the other of the two variations depending upon which gives the most information. It is called the either-or hypothesis ($H_{A/B}$).

In order to evaluate these hypotheses, it was first necessary to obtain the P_A and P_B values for each trial. This was accomplished by two methods. By one method (true proportions) it was assumed that the best estimate of P_A, for example, was the objective probability of the particular A variation being an X. Although it was evident that this could not be the case in the early trials of the experiment, it was considered worth investigating since the objective probability might eventuate in a good approximation of P_A.

The second method (single stimulus) was based upon the results obtained from the Stimulus A and Stimulus B groups. The proportion of subjects answering X was determined for each of the 420 trials for both these groups. These proportions were used to estimate P_A and P_B. For example, if 12 of the 18 subjects of the Stimulus B Group perceived the picture as an X, then P_B was given the value of 12/18 = .666 for that trial. This allowed the value of P_A or P_B for any particular variation to change in the course of the experiment in accordance with the empirical evidence.

It is of interest at this point to examine the change in P_A, as calculated from the single stimulus method, as the experiment progressed. This is shown with smoothed curves in Figure 6.

At the beginning of testing the P_A values of the stimulus variations were, as would be expected, at or near .5. As testing progressed, however, they underwent the changes exemplified by Figure 6. It is evident that the values did not tend to become approximations of the true probability function. Rather, they approached a function which can be defined as the "best strategy." The "best strategy" function is also shown in the figure.

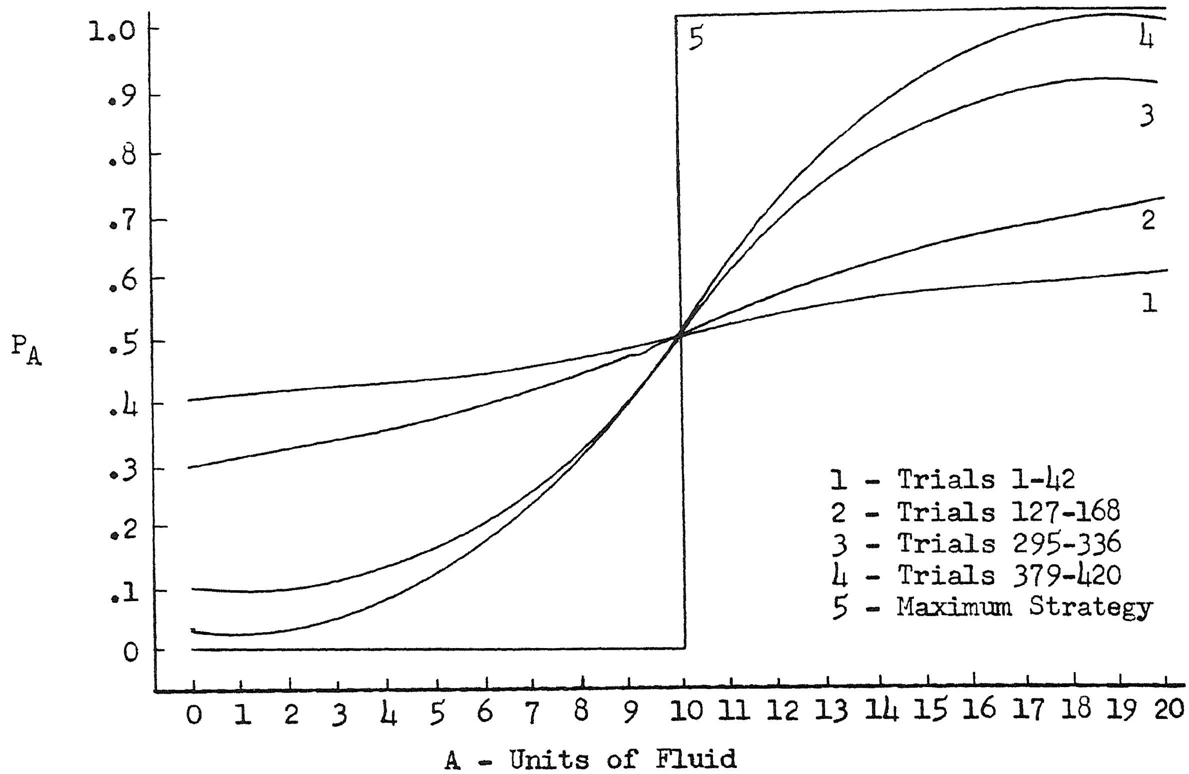


Figure 6

P_A as a Function of the Variation in A

Table 50

Correlations between Proportion of Subjects Responding "X" and the Proportion Predicted to Respond "X" on the Basis of the Various Hypotheses Tested

	True Proportions				Single Stimuli			
	T* 1-105	T* 106-210	T* 211-315	T* 316-420	T* 1-105	T* 106-210	T* 211-315	T* 316-420
H_A	.218	.845	.810	.839	.545	.856	.781	.826
H_B	.168	.657	.696	.721	.530	.699	.649	.778
H_{A+B}	.306	.817	.907	.915	.670	.888	.942	.971
H_{AB}	.355	.875	.945	.943	.631	.694	.924	.945
$H_{A/B}$	.376	.850	.868	.899	.645	.830	.830	.943
$H_{\sqrt{A+B}}$	.227	.895	.851	.770	.611	.746	.820	.757

* T - Trials

It is defined as the function which would yield the greatest possible amount of reinforcement. It is significant, however, that the P_A values tended to stabilize on points somewhat removed from this line of "best strategy."

P_B produced a similar variation in time.

Having obtained estimates of P_A and P_B it was then possible to predict, on the basis of the hypothesis described above, the proportion of the Two Stimulus Problem subjects who would perceive a picture as an X. Predictions were generated for each trial using each hypothesis and using the estimates of P_A and P_B obtained from the single stimuli procedure. Other predictions were obtained for each hypothesis using the objective probabilities (true proportions). Correlations were then calculated between the prediction and the actual outcome.

The correlations obtained are shown in Table 50. Separate correlations have been calculated for each block of 105 trials.

By this method of evaluation the estimates of P_A and P_B based on the single stimulus procedure are better than those based on the true proportions. This is especially true of the early trials of testing.

It is also evident that H_A and H_B are inferior to the hypotheses which involve a combination of elements. Of these latter hypotheses, $H_{\sqrt{A+B}}$ is obviously inferior.

It is a more difficult matter to distinguish between the remaining hypotheses. H_{A+B} is generally the highest for the single stimuli method followed by H_{AB}. With the true proportion method, H_{AB} yields the highest correlation.

The significance of the difference between some of these correlations is shown in Table 51. Most important is the fact that for the single stimulus method, H_{A+B} is significantly superior to all other hypotheses in the last block of trials.

DISCUSSION

Even had testing been extended indefinitely, it does not appear that the subjects on the Stimulus A and Stimulus B Problems would have ever adapted the "strategy" which would yield the maximum amount of reinforcement. This strategy can be visualized from the curves in Figure 6. A subject would achieve the maximum amount of reinforcement by always calling any stimulus variation an X if its probability of being an X was in excess of .50, and always calling the variation a Y if its probability of being Y was less than .50. The Single Stimulus groups tended toward this strategy but never completely accepted it.

It is interesting to speculate on why human judgement evolves in such a manner that a "perfect solution" to these simple problems is not obtained. It is doubtful, in the present instance at least, if the difficulty is due to an inability to make the necessary stimulus discriminations. Even when subjects appear to know that a variation is

probably an X, they continue to make occasional Y responses. One is tempted to conclude that the human perceptual process is a very faulty computer.

Table 51

Significance of the Difference between some of the Correlations shown in Table 50

	True Proportions				Single Stimuli			
	Trials				Trials			
Comparison	1-105	106-210	211-315	316-420	1-105	106-210	211-315	316-420
H_A vs H_B	NS	$P<.01$	NS	$P<.05$	NS	$P<.01$	NS	NS
H_A vs $H_{\sqrt{A+B}}$	NS	NS	NS	NS	NS	$P<.05$	NS	NS
H_B vs $H_{\sqrt{A+B}}$	NS	$P<.01$	$P<.01$	NS	NS	NS	$P<.01$	NS
H_{A+B} vs H_{AB}	NS	NS	NS	NS	NS	$P<.01$	NS	$P<.05$
H_{A+B} vs $H_{A/B}$	NS	NS	NS	NS	NS	NS	$P<.01$	$P<.05$
H_{AB} vs $H_{A/B}$	NS	NS	$P<.01$	$P<.05$	NS	$P<.05$	$P<.01$	NS

On the other hand, it is possible to consider it a very sophisticated computer. Unlike man-made computers, it has not been programmed for a particular problem. It has not been told, for example, that variation A and only variation A is relevant to the solution. Moreover, the brain evolved in an environment where outcomes as a function of stimulus variables were not constant, but rather they underwent modification in time.

If the organism were to achieve a perfect solution (i.e., responding 100% to probabilities in excess of .50 and 0% to probabilities of less than .50) with respect to one stimulus variable or combination of variables, this would completely desensitize it to any new stimulus variables which might become equally useful. Thus stimulus synthesis could not be obtained.

The human computer appears to settle for a solution that lies somewhere between the "perfect solution" and the objective probability value. In so soing, it loses in precision but gains greatly in adaptability.

The poor showing made by the $H_{\sqrt{A+B}}$ analysis is surprising in view of the results of the previous Section. This may be due in part to the fact that in the earlier study essentially twenty times as much information was given on each trial as was given in the present study.

If this is true, it could be inferred that if the present procedure were continued far beyond the 420 trials given, that the $H_{\sqrt{A+B}}$ would gain in predictive power.

SUMMARY

Three groups of subjects were tested by being shown a series of 420 pictures. These groups were required to discriminate between the pictures which were X's and which were Y's. Each picture showed a glass graduate (which varied in the amount of fluid it contained) and a marble board (which varied in the number of marbles present). As the graduate gained in fluid or the marble board showed a decrease in marbles, the probability that the picture was an X increased. X's and Y's were differentiated on the basis of these variations. One group of subjects saw only the glass graduate. A second group saw only the marble board. A third group saw both stimuli.

A number of analyses were conducted in an effort to determine the basis for the decisions made by the group seeing both stimuli. It was concluded that for this procedure, and for the number of trials given, stimulus synthesis is best described as a process in which the response strengths elicited by the individual stimulus determinants are summed. There is also evidence to support the conclusion that a disproportional emphasis is placed upon the stimuli generating very high or very low response tendencies.

SECTION 12

STIMULUS SYNTHESIS INVOLVING FOUR STIMULUS VARIATIONS

Procedurally this study was similar to the experiment described in Section 11. It differs from the previous problem by exercising four stimulus variables rather than two.

The purpose of this experiment was threefold. First the intention was to extend the earlier study by determining how well the formulas held up when applied to four stimuli. With four relevant stimuli the multiplicative hypothesis, for example, would take the following form:

$$P_X = \frac{P_A \, P_B \, P_C \, P_D}{P_A \, P_B \, P_C \, P_D + (1-P_A)(1-P_B)(1-P_C)(1-P_D)}$$

The other formulas would undergo similar revision.

The second purpose was to vary the absolute probability parameters associated with the stimulus variations. In nature, the distribution of a class of things with respect to a sensory continuum can take many forms. Admittedly, classes are not the product of nature. They are the products of man's tendency to organize or to conceptualize upon nature. This only illustrates, however, that man is capable of forming concepts based upon stimulus relationships that are almost amorphous in their probability distributions. In the present study, it was conceived that greater generality would be obtained if at least one distribution of X, with respect to a stimulus variation, was less systematic than the distributions previously used.

The final purpose was to test retarded subjects on these problems, as well as normal subjects.

Unfortunately, the results did not go along with the plans. First, it was found that retarded subjects could not learn to discriminate between X and Y by the method used. Second, it was found that even normals had great difficulty in using two of the stimulus cues provided. This was not due to the indiscriminability of stimulus variation, but rather to the overlapping of the X and Y distributions with respect to these cues.

METHOD

Subjects: In all, 331 subjects were tested. Of these, 95 were high school students, 62 were from junior high school, 132 were from grade school, and the remaining 42 were mentally retarded.

Apparatus and Stimuli: As in the earlier study a slide projector, a screen, and a group of 35mm color transparencies were used. For 420 of the

pictures, four objects were shown. These were (A) a jigsaw puzzle which varied in degree of completion, (B) a marble board which varied in the number of marbles it held, (C) a clock which varied in the time of day shown, and (D) a glass cylinder which varied in the amount of fluid it contained. Two hundred and ten of these pictures formed the population of X's and the remaining formed the population of Y's. X's could be discriminated from Y's on the basis of variation shown by the cues.

The probability of the picture being an X when only one piece of the puzzle was present was zero. As the puzzle became more complete, however, the probability of it being an X increased up to a maximum value (when 10 or 11 pieces were present) after which the probability decreased again to zero. This function is shown in Figure 7. Since each picture was either an X or a Y, the probability of a presentation being a Y on the basis of the variation in A was greatest at extremes of completion and incompletion.

The probability of a picture being an X on the basis of the B variation decreased as the number of marbles increased. This is shown in Figure 8.

The probability of a picture being an X on the basis of the time of day shown on the clock (C), increased as the minute hand moved forward. This function is presented in Figure 9. Notice, however, that the slope of the curve is gradual. The probability of X never exceeded .75 or fell below .25 for any value of C.

Finally, the probability of a picture being an X as determined by the amount of fluid in the glass cylinder is shown in Figure 10. Here the slope is very abrupt. When either one, two, three, four, or five units of fluid appeared, the picture was always a Y. At the opposite extreme, the picture was always an X.

Except for the above restriction, the association of the A, B, C, and D variation was random in the pictures for both the X and Y populations.

In addition to the 420 pictures in which the four objects were displayed, photographs were also obtained of each object alone. Twenty-one pictures were taken of A in each of its variation, twenty-one pictures of B, etc. These were used in connection with the single stimulus problems.

Procedure: Five stimulus treatment conditions were used. The Composite Stimuli Group was shown each of the 420 pictures displaying all four objects. These were presented in random order. On each presentation they were required to judge whether the picture was an X or a Y. After they had registered their response, they were told the correct answer. The pencil passing technique described in Section 10 was used to prevent responses from being recorded after the answers were given.

Except for the retarded children (who were given individual attention) group testing was used. The retarded subjects were not required to make X and Y responses. Instead, they indicated a preference between two squares and were told when they were right and wrong.

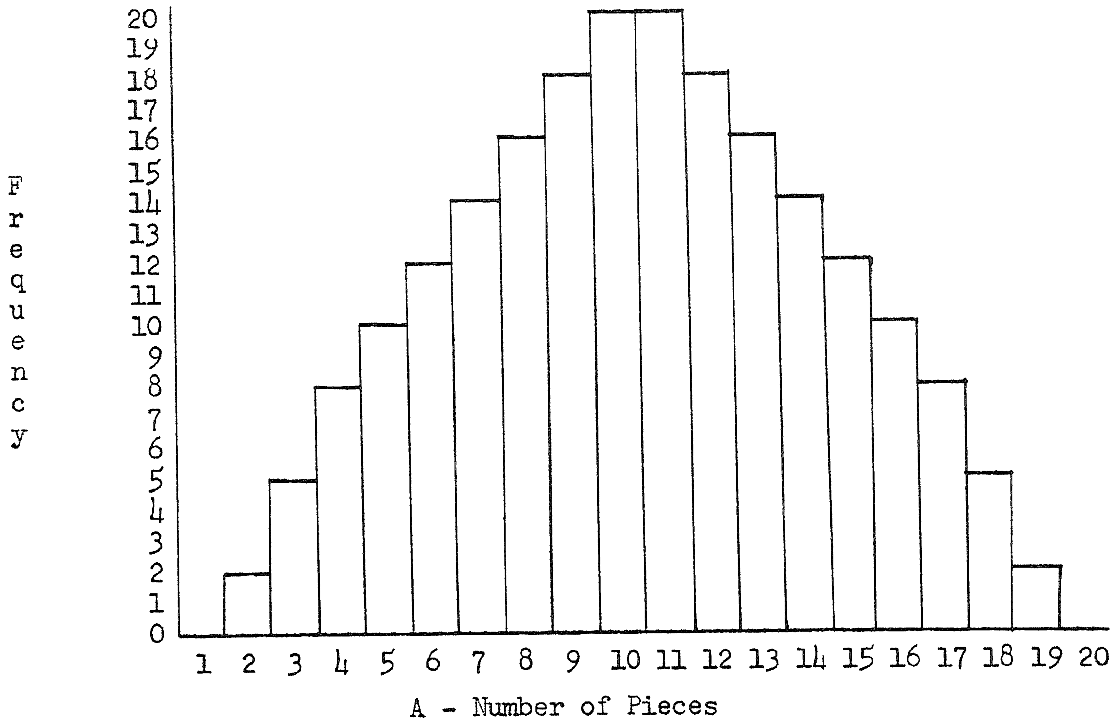


Figure 7

Frequency of X as a Function of the Variation in A

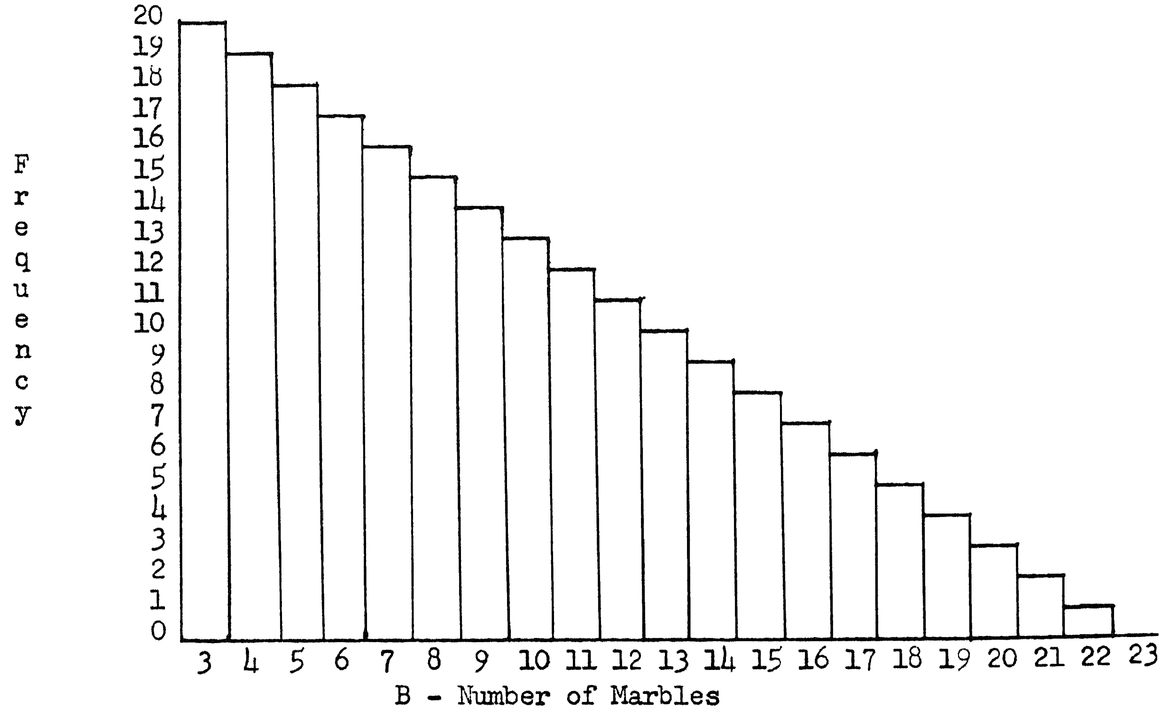


Figure 8

Frequency of X as a Function of the Variation in B

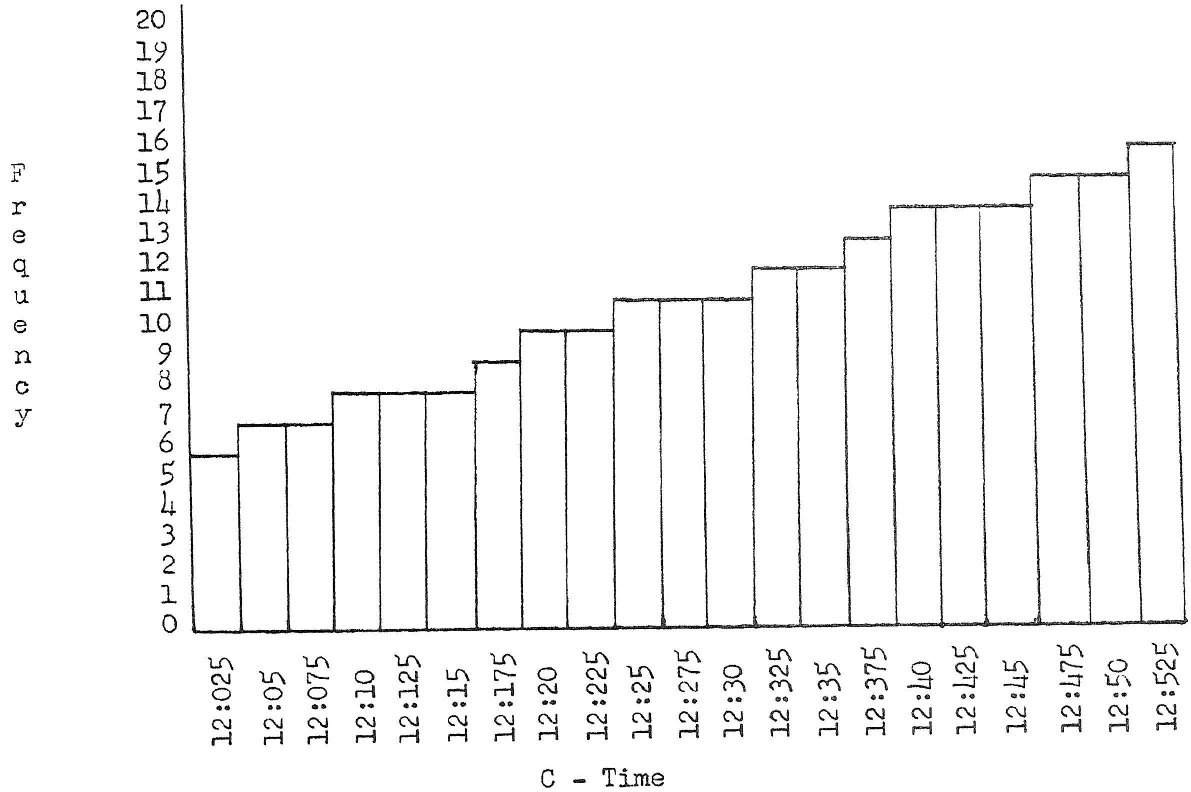


Figure 9

Frequency of X as a Function of the Variation in C

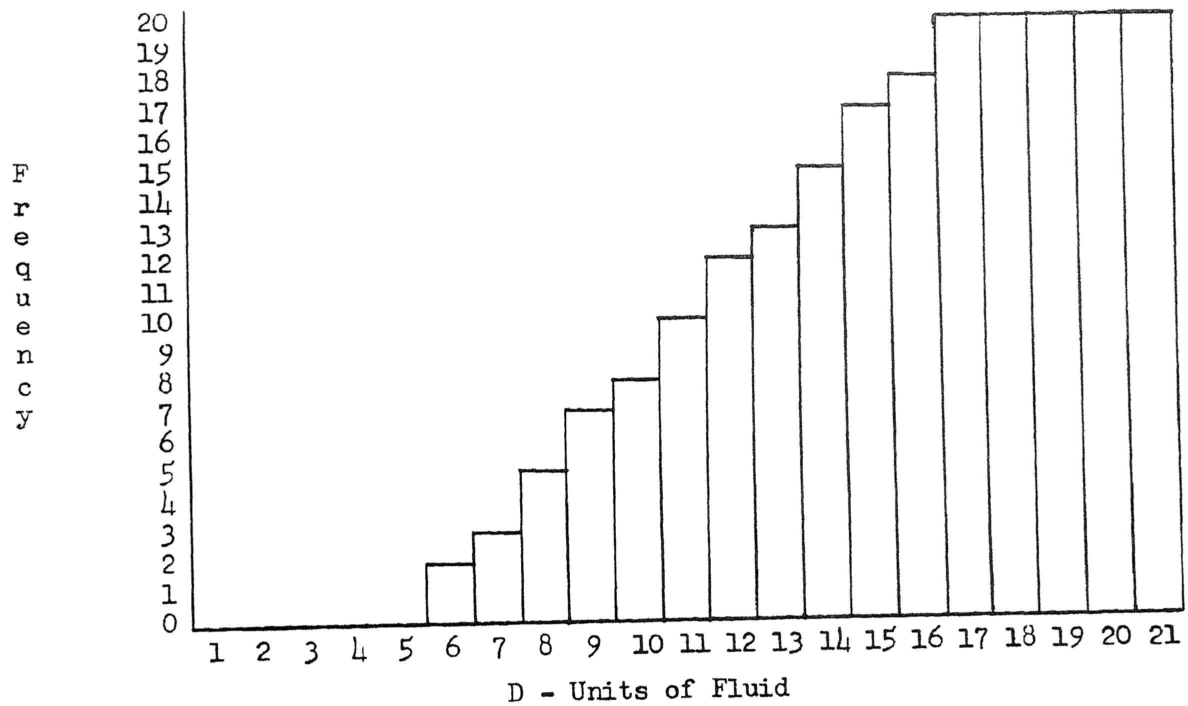


Figure 10

Frequency of X as a Function of the Variation in D

The Stimulus A Group was also given 420 presentations to judge, but they saw only the pictures in which A was alone. The sequence of the A variation was identical with the sequence obtained for A in the composite stimuli. The Stimulus B Group saw only the B variation while the Stimulus C and Stimulus D Groups saw only C and D respectively.

The retarded subjects were all assigned to the composite stimulus treatment. It was originally planned to use other retarded subjects on the single stimulus problems, but the fact that they were unable to solve the Composite Stimulus Problem pre-empted this requirement.

RESULTS

Considering first the single stimulus groups, it was found that the Stimulus A and the Stimulus C problems were extremely difficult to solve.

The B and D problems were comparatively simple. In general, the response tendencies for B and D resemble those shown in Figure 6. Beginning with a non-differentiating function, the curve gradually shifted and finally came to rest between the absolute probability curve and the "best strategy" function. To some extent, the response tendencies of the A and C Stimulus problems evolved in a similar manner, but the inter-trial variations were greater.

The differences between subject groups on the single stimulus problems favored the high school subjects but were not statistically significant. They are, for this reason, combined for further analysis.

The Composite Stimuli Problem was given to high school, junior high school, grade school, and retarded subjects. The retarded subjects could not solve the problem. Their responses remained at chance level throughout testing. This should not be interpreted as meaning that these subjects cannot discriminate between multi-variate stimuli. The fault lies with the method used. It is believed that had the presentation involved two pictures with the requirement that the subject select the "X-iest" (the one with the greater probability of candy reinforcement), retarded subjects would have solved the problem. The data on retarded children was not subjected to further analysis.

In Table 52, an analysis of variance is run on each block of 42 trials comparing correct responses for high school, junior high school, and grade school groups. The differences are significant on only three blocks of trials (253-294, 295-336, and 337-378). The means for these blocks are shown in Table 53 and the significance of the difference between individual means is presented in Table 54.

As in the Section 11 analysis, P_A (tending to perceive a picture as an X based upon the A variation), P_B, P_C, and P_D were estimated for each trial from the results of the single stimulus problems. These results were applied to the linear, multiplicative, square root, and the either/or formulas. All possible stimulus variations were considered for each formula. Each formula had the following stimulus combinations: AB, AC, AD, BC, BD, CD, ABC, ABD, ACD, BCD, and ABCD. Correlations were

Table 52

Analysis of Variance based on Correct Responses
by High School, Junior High,
and Grade School Groups

Trials 1-42

	SS	df	MS	F	P
Schools	14.39	2	7.20	-	
Within	2512.62	142	17.69		
Total	2527.01	144			

Trials 43-84

	SS	df	MS	F	P
Schools	121.98	2	60.99	1.86	
Within	4661.77	142	32.83		
Total	4783.75	144			

Trials 85-126

	SS	df	MS	F	P
Schools	112.55	2	56.28	1.78	
Within	4495.08	142	31.66		
Total	4607.63	144			

Trials 127-168

	SS	df	MS	F	P
Schools	104.09	2	52.05	-	
Within	8253.40	142	58.12		
Total	8357.49	144			

Trials 169-210

	SS	df	MS	F	P
Schools	147.64	2	73.82	1.50	
Within	6976.95	142	49.13		
Total	7124.59	144			

Trials 211-252

	SS	df	MS	F	P
Schools	185.34	2	92.67	2.23	
Within	5898.00	142	41.54		
Total	6083.34	144			

Trials 253-294

	SS	df	MS	F	P
Schools	190.79	2	95.40	4.38	<.05
Within	3090.56	142	21.76		
Total	3281.35	144			

Trials 295-336

	SS	df	MS	F	P
Schools	732.22	2	366.11	26.82	<.01
Within	1938.84	142	13.65		
Total	2671.06	144			

Trials 337-378

	SS	df	MS	F	P
Schools	297.80	2	148.90	4.89	<.01
Within	4328.03	142	30.48		
Total	4625.83	144			

Trials 379-420

	SS	df	MS	F	P
Schools	95.63	2	47.82	1.97	
Within	3444.81	142	24.26		
Total	3540.44	144			

Table 53

Mean Correct Responses by Subject Groups for the 253-294, 295-336, and 337-378 Blocks of Trials

	School		
Blocks	High School	Junior High	Grade School
253-294	31.10	30.03	28.04
295-336	31.10	29.67	25.40
337-378	36.68	35.09	32.78

Table 54

_t_ Scores Resulting from Subject Group Comparisons on Blocks 253-294, 295-336, and 337-378

Blocks	Comparison*	t	P
253-294	HS - GS	2.89	<.01
	JH - GS	2.12	<.05
	HS - JH	1.16	NS
295-336	HS - GS	7.81	<.01
	JH - GS	5.55	<.01
	HS - JH	1.86	NS
337-378	HS - GS	3.22	<.01
	JH - GS	2.03	<.05
	HS - JH	1.46	NS

* HS - High School
JH - Junior High
GS - Grade School

obtained between the proportion of the composite stimuli subjects perceiving each picture as an X and the proportion predicted by the formulas. In essence, the correlations are between the scores on the Composite Stimulus Problem and the scores (subject to the transformation of the formulas) on the single stimulus problems. In these analyses it became evident that the Composite Stimulus Problem was solved without reference to the A and C stimulus variations. Only B and D were used in any systematic way and the B variation was obviously of secondary importance to the D variation.

It was also determined that although some of the correct response scores between subject groups differed significantly (as seen in Table 54) the correlations obtained for these groups did not differ to a significant degree. Consequently, all groups except the retarded subjects were combined for the correlations shown in Table 55. In this table only the hypotheses based upon the B and D combinations are shown.

Table 55

Correlations between Predicted and Actual Proportion of Subjects Responding X

	Trials			
Hypothesis	1-105	106-210	211-315	316-420
H_B	.351	.413	.431	.479
H_D	.789	.921	.969	.971
H_{B+D}	.609	.787	.812	.916
H_{BD}	.561	.804	.874	.920
$H_{B/D}$	.522	.835	.908	.900
$H_{\sqrt{B+D}}$	.599	.738	.777	.825
H_A	.740	.443	.331	.380
H_C	.284	.318	.296	.300

It can be seen that the subjects based their judgments almost entirely on the D variation. Indeed the hypothesis H_D achieves a higher correlation with the results than do any of the combinational hypotheses. This should bot be interpreted to mean that stimulus synthesis was not occurring. It does mean, however, that D was being given a great deal of emphasis.

The significance of the difference between correlations is presented in Table 56. Generally H_D is shown to be the best predictor. Among the combinational formulas, $H_{\sqrt{B+D}}$ proved the poorest predictor.

DISCUSSION

In a sense, this study was a failure. Its intent was to shed light on the problem of how normal and retarded children integrate stimulus cues into unitary percepts. Instead, it uncovered some conditions under which stimulus synthesis either does not occur or cannot be detected. It did, however, provide some information of use in future experimentation. Some of these are listed below.

1. The method cannot be used with retarded subjects. These children simply did not get the "connection" between the pictures and the response categories used. It is believed that the problem can be given to this type of subject if it is presented as a discrimination situation in which two pictures are simultaneously presented and the subject is required to choose between them. The type of apparatus described in Section 7 should be useful for this purpose.

2. The stimulus distributions with respect to X were not appropriate. This is especially true of Stimulus D. Had it not been so useful a cue, the subjects would have had to rely on the more difficult discriminations. When certain stimulus variations are invariantly the property of a particular response class, subjects tend to base their judgments almost entirely on these variations. This is intelligent behavior, but it tells us nothing about stimulus synthesis. It may also be that Stimuli A and C were too difficult, but this could not be ascertained in the presence of D.

3. A random association of stimulus variations tends to produce many examples of very high or very low composite probability values. This makes it difficult to distinguish between the hypotheses tested. A better method would be to control the absolute probabilities so that they distribute themselves in a more normal fashion. This can be accomplished by correlating the stimulus variations. This procedure presents an interesting question. Since highly correlated stimuli are also redundant to one another, will learning occur as rapidly as when the stimuli are not correlated?

On the more positive side, it was found that $H_{\sqrt{B+D}}$ was inferior to the other composite hypotheses as a predictor. This is in agreement with the results in Section 11, but is not altogether in agreement with the results of Section 10. It will require further investigation and analysis to clear up this apparent discrepancy.

SUMMARY

Three hundred and thirty-one normal and retarded subjects were assigned to five treatment conditions. Stimulus A Group was shown pictures of a jig-saw puzzle which varied in its degree of completeness. Stimulus B Group saw

Table 56

Significance of the Difference Between the Correlations Presented in Table 55

Trials 1-105

Hypotheses	H_D	H_{B+D}	H_{BD}	$H_{B/D}$	$H\sqrt{B+D}$
H_B	<.01	<.05	NS	NS	NS
H_D		NS	<.01	<.01	<.01
H_{B+D}			NS	NS	NS
H_{BD}				NS	NS
$H_{B/D}$					NS

Trials 106-210

Hypotheses	H_D	H_{B+D}	H_{BD}	$H_{B/D}$	$H\sqrt{B+D}$
H_B	<.01	<.01	<.01	<.01	<.01
H_D		<.01	<.01	<.01	<.01
H_{B+D}			NS	NS	NS
H_{BD}				NS	NS
$H_{B/D}$					NS

Trials 211-315

Hypotheses	H_D	H_{B+D}	H_{BD}	$H_{B/D}$	$H\sqrt{B+D}$
H_B	<.01	<.01	<.01	<.01	<.01
H_D		<.01	<.01	<.01	<.01
H_{B+D}			NS	<.01	NS
H_{BD}				NS	<.05
$H_{B/D}$					<.01

Trials 316-420

Hypotheses	H_D	H_{B+D}	H_{BD}	$H_{B/D}$	$H\sqrt{B+D}$
H_B	<.01	<.01	<.01	<.01	<.01
H_D		<.01	<.01	<.01	<.01
H_{B+D}			NS	NS	<.01
H_{BD}				NS	<.01
$H_{B/D}$					<.05

pictures of a marble board in which the number of marbles varied. Stimulus C Group was given pictures of a clock. The time of day varied between pictures. Stimulus D Group was presented with pictures of a glass cylinder in which the amount of fluid contained varied. The Composite Stimuli Group saw pictures in which puzzle, marble board, clock, and cylinder were simultaneously present and each varied in the same manner as they did in the single stimulus treatments.

The subjects were required to classify each picture as an X or a Y. After each trial, they were told if their response was correct or not.

The results generated by the single stimulus groups were combined in various ways to predict the outcome of the Composite Stimulus Treatment.

It was found that the retarded subjects could not solve the Composite Stimulus Problem and that the normal subjects based their judgments primarily on the variation of Stimulus D.

SECTION 13

GENERAL SUMMARY

The experimental findings obtained under this contract are described as eleven separate studies in the report. In Section 1, a review of the literature on concept formation and a general discussion is presented. An account of the experiments is given in Sections 2 through 12.

The purpose of the investigation was to study the development of concept formation as it is affected by (1) the type of subject used (normal or retarded), (2) problem (conceptual organization required for solution), (3) stimulus differences, (4) response required (manual or verbal), and (5) reinforcement schedule. The procedure varied between experiments. In some studies, a double stimulus presentation was made while in others the single stimulus procedure was adopted. In all of the studies, the subjects were restricted to a choice between two responses.

Some of the experiments were not large with respect to the sample of subjects tested and some did not employ retarded subjects. The policy usually followed was first to investigate the effect of a limited number of variables on normal subjects. Second, those variables which appeared to be of interest were later combined into a study involving retarded as well as normal subjects.

In Section 2 entitled, "Effect of Type of Response on the Solution of the Similarity-Difference Problem," normal grade school children served as subjects. All subjects received the Similarity-Difference Problem, but the stimulus and response conditions varied between groups. Half of the subjects were shown humanoid stimuli (pictures of humans who varied in sex, age, race, and position). For 12 of these subjects, sex (male vs female) was the relevant concept and the other stimulus variations were irrelevant. For the remaining 12 subjects presented the humanoid stimuli, the age difference (child vs adult) was relevant.

All other subjects received the geometrical stimuli (pictures of parallel lines which differed in color, thickness, orientation, and background). For 12 of these subjects, the orientation of the lines was relevant and for the remaining 12, the color of the lines was relevant.

Each stimulus treatment group was further divided into three subgroups with respect to type of response required. One group (Direct Manual) was directed to respond by pressing switches located below the stimulus windows. The second group (Indirect Manual) responded by pressing switches located between the stimulus windows. The third group (Verbal) responded verbally.

The differences between stimulus treatments were not significant. The differences between response treatments were statistically significant favoring the verbal treatment over the indirect manual treatment and the indirect manual treatment over the direct manual treatment. The interaction was not significant.

These results are interpreted as indicating that the development of a concept is dependent upon the response required. Verbal responses, in particular, facilitate relational learning. Direct Manual responses, which tend to break the configuration into its stimulus components, retard learning.

Section 3 which is entitled, "Comparison of Four Grade Levels on the Similarity-Difference Problem," was a continuation of the first study. The development of the Similarity-Difference concept was shown to be related to the subject's age. The verbal response was slightly superior to the indirect manual response. It was also witnessed in this study that the concept of race was easier to obtain than the concept of sex and that the concept of age was the most difficult to obtain of the concepts tested.

In Section 4, "Concept Reversal Under Two Problem Conditions," 16 groups were delineated with respect to treatment. There were three stages to the experiment. In the first stage, half of the groups received the Similarity-Difference Problem while the other half received the Double-Successive Problem. In the second stage, a new problem relationship was conjoined to the original problem. For half of the groups, the new problem was conceptually identical (although the stimuli differred) with the old problem. For the other groups, the new problem involved the opposite relationship. In addition to the stimuli associated with the conjoined problems of the second stage, irrelevant stimuli were also introduced.

In the third stage, all subjects received a problem that was conceptually identical with the conjoined problem of the second stage. For half of the groups, the stimuli of relevance in the second stage was also relevant in the third stage. For the other groups, the irrelevant stimuli of the second stage became the relevant stimuli in the third stage.

Groups were balanced off with respect to stimulus differences.

The hypothesis being tested was that subjects apprehend all stimulus-response relationships which can be organized within the conceptual model being developed, but fail to apprehend other available solutions to the problem. The results failed to confirm the hypothesis. The evidence pointed to the coexistance of the two conceptual processes.

In Section 5 which carries the title, "Comparison of Normal and Retarded Children on the Similarity-Difference Problem," an experiment is described in which the mental age of the subjects, the stimulus of relevance, the number of stimulus variables, the types of response, and the reinforcement procedures were varied. This study was done in three stages. One hundred percent reinforcement was given in the first stage and 50% in the second stage. In the third stage, there was a reversal in the stimulus of relevance. Concurrent with this reversal, half of the subjects were raised back to 100% reinforcement and the other half were reinforced at the 50% rate.

It was hypothesized that the 50% reinforcement rate would be superior because these subjects would drop below their expected reinforcement level. The 100% reinforcement subjects, on the other hand, would obtain their

expected level even by responding to the improper stimuli. The results attested to the superiority of the normal over the retarded subjects, of the verbal over the manual response, and of the race stimulus variable over the age and sex stimulus variables. The reinforcement variable was not significant. No interactions of significance were found.

Section 6, "Comparison of Two Reinforcement Schedules on Problem Reversal," is a further test of the hypothesis that the efficacy of reinforcement is a function of its value relative to the subjects' expectation rather than its absolute value. Only normal fifth grade subjects were used in this study.

The stimuli consisted of three lights set in a panel. With respect to illumination and non-illumination, there are eight possible combinations of three lights. It was necessary for the subject to learn the proper response (pushing a toggle switch either to the left or to the right) to each combination. The combinations associated with a particular response were selected so that the solution could be obtained on some conceptual basis such as "when the top light is on, the left response is correct; when it is off, the right response is correct."

In the first stage of the experiment, 100% reinforcement was given all subjects. In the second stage, the subjects were transferred to 50% reinforcement. In the third stage, the problem was changed for all subjects without their being told of this. In the third stage, half of the subjects went back to 100% reinforcement and the other half remained on 50% reinforcement.

In this study, the 100% reinforcement group was superior to the partial reinforcement group. In view of the negative results between the reinforcement treatments of the previous problem, this outcome suggests that the predicted effect exists, but is not strong enough to overcome the loss of information occasioned with the 50% reinforcement schedule. Subjects can learn to respond to new stimuli under 50% reinforcement, but have extreme difficulty learning an entirely new concept under this schedule.

Section 7 is entitled, "The Effect of Varying the Probability and the Amount of Reinforcement in Normal and Retarded Subjects." In this experiment, the subjects learned an order of preference between ten stimulus patterns. The subject groups were divided with respect to reinforcement procedures. For the probability treatment, Stimulus A was reinforced 90% of the time, Stimulus B 80% of the time, Stimulus C 70% of the time, etc. For this treatment, reinforcement always consisted of 10 small pieces of candy. For the amount treatment, Stimulus A was always reinforced, but the amount of reinforcement consisted of nine pieces of candy. Similarly, Stimulus B was always reinforced, but the reinforcement was eight pieces of candy, etc. Over a period of trials, a consistent choice would yield the same amount of reinforcement whether by the probability or the amount treatments.

A correct choice, for both treatments, consisted of choosing the stimulus having the greater probability or the greater amount of reinforcement. It was found that fifth and sixth grade subjects were superior to first and second graders and that retarded subjects were inferior to

either of the normal groups. It was also witnessed that the amount treatment tended to facilitate learning in the normal subjects as compared to the probability treatment. For the retarded subjects, however, this relationship was reversed. The probability treatment was superior to the amount treatment for these subjects.

In Section 8, "Effect of Partial Reinforcement on Normal and Retarded Subjects," the hypothesis was tested that partial reinforcement procedures would affect normal and retarded children differentially. Two treatments were involved. The 100% reinforcement treatment subjects were given candy for each correct response. Along with the candy, a red poker chip (secondary reinforcement) was given. A blue poker chip was given for an incorrect choice. For the 50% reinforcement treatment, the red poker chip was given on each correct trial and the blue poker chip was given for each incorrect trial, however, candy was given on only half of the times that the red poker chip was given.

The normal subjects were superior to the retarded children. The difference in reinforcement favored the 100% treatment, but this was not significant. The interaction (upon which the hypothesis was to be validated) was also without significance. The hypothesis was not confirmed by the data.

Section 9, "Concept Formation in the Solution of the Three Light Problem," examines the hypothesis that the solution to a problem is facilitated if it can be easily verbalized. Using the three light situation, problems having easily verbalized solutions were compared with problems whose solutions were difficult to verbalize. The results tended to support the hypothesis, but the effect was not as strong as expected. Many of the problems appeared to have been solved on the basis of two or more limited concepts, rather than one unifying concept.

In Section 10, "Comparison of Three Methods of Reinforcement Quantification," the act of stimulus synthesis is examined. A population of 420 pictures was produced for this study. In each picture, two objects (A and B) were shown. A was a glass cylinder which varied from picture to picture in the amount of fluid contained. B was a marble board which varied in the number of marbles present. Half of the pictures were classified as X's and the remaining were classified as Y's. The probability that a picture was an X, depended upon the amount of fluid in A and the number of marbles in B. As the fluid increased or as the number of marbles decreased, the probability of the picture being an X increased.

The procedure was based upon the assumption that a reinforcement schedule which is quantitatively proportional to the mental process of stimulus integration, will produce better learning than a reinforcement schedule that is not proportional. Three hypotheses concerning the process of stimulus synthesis were compared. The linear hypothesis states that integration is proportional to the respective probabilities of the separate objects being X's. The multiplicative hypothesis states that subjects synthesize primarily on the basis of the extreme probability values. The square root hypothesis states that integration involves a disproportionate amount of attention to the central probability values.

Three treatments were exercised. One group obtained a reinforcement schedule consistent with the linear hypothesis. A second group was given reinforcement consistent with the multiplicative hypothesis and the third group received reinforcement based upon the square root hypothesis.

The results indicated that high school subjects were initially facilitated by the multiplicative reinforcement schedule, but later obtained somewhat better scores using the linear schedule. Grade school subjects appeared to do better with the square root schedule of reinforcement.

In Section 11, "Perceptual Synthesis Involving Two Stimulus Variations," an experiment is reported in which the stimuli were the same as those described for Section 10. Three treatment groups were tested. Stimulus A Group saw only the A object. Stimulus B Group saw only the B object. The Two Stimulus Group saw both objects. Subjects were required to guess whether each picture was an X or a Y.

The analysis consisted of obtaining correlations between the proportion of the two stimulus problem subjects perceiving a picture as an X and the proportion predicted on the basis of the various hypotheses. The probability values of X on the basis of A and B were derived from the data on the stimulus A and stimulus B subjects respectively. The predictions based upon the multiplicative hypothesis were generally the most accurate. The square root hypothesis was the weakest in predictive power.

The last experiment of the series is reported in Section 12. Entitled, "Stimulus Synthesis Involving Four Stimulus Variations," it represents a logical extension to the study described in Section 11. It differed from the earlier study in that: (1) four objects rather than two were involved, (2) retarded subjects as well as normals were used, and (3) the distribution of X as a function of the variations of A, B, C, and D were highly dissimilar. The experiment failed in its primary purpose. It did show, however, that mentally retarded subjects could not solve the problem under the test conditions used. It also indicated that subjects tend to use high probability functions to the exclusion of the less reliable cues.

REFERENCES

1. Bitterman, M. E. Spence on the problem of patterning. Psychol. Rev., 60, 1953, 123-126.

2. Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: John Wiley and Sons, 1956, 103-112.

3. Denny, M. R. The role of secondary reinforcement in a partial reinforcement learning situation. J. exp. Psychol., 1946, 36, 373-89.

4. Elam, C. B., Tyler, D. W., & Bitterman, M. E. A further study of secondary reinforcement and the discrimination hypothesis. J. comp. physiol. Psychol., 1954, 47, 381-384.

5. Fields, P. E. Studies in concept formation: I. The development of the concept of trangularity by the white rat. Comp. Pshchol., Monogr., 1932, 9, No. 2.

6. Hanfmann, E., & Kasanin, J. A method for the study of concept formation. J. Psychol., 1937, 3, 521-40.

7. Heidbreder, E. The attainment of concepts: I. Terminology and methodoloty. J. gen. Psychol., 1946a, 35, 173-89.

8. Heidbreder, E. The attainment of concepts: II. The problem. J. gen. Psychol., 1946b, 35, 191-223.

9. Heidbreder, E. The attainment of concepts: III. The process. J. gen. Psychol., 1947, 24, 93-138.

10. Hull, C. L. Quantitative aspects of the evolution of concepts. Psychol. Monogr., 1920, No. 123.

11. Hull, C. L. The concept of the habit-family-hierarchy and maze learning: Part I. Psychol Rev., 1934a, 41, 33-54.

12. Humphreys, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. J. exp. Psychol., 1939a, 25, 141-58.

13. Krechevsky, I. A study of the continuity of the problem-solving process. Psychol. Rev., 1938a, 45, 107-33.

14. Kreezer, G., & Dallenbach, K. M. Learning the relation of opposition. Amer. J. Psychol., 1929, 41, 432-41.

15. Kohler, W. The mentality of apes. New York: Harcourt, Brace, 1925.

16. Lashley, K. S. An examination of the 'continuity theory' as applied to discrimination learning. J. gen. Psychol., 1942, 26, 241-65.

17. Osgood, C. E. Method and theory in experimental psychology. New York: Oxford University Press, 1953, 666-667.

18. Reed, H. B. Factors influencing the learning and retention of concepts: I. The influence of set. J. exp. Psychol., 1946a, 36, 71-87.

19. Reed, H. B. The learning and retention of concepts: II. The influence of length of series. III. The origin of concepts. J. exp. Psychol., 1946b, 36, 166-79.

20. Reed, H. B. The learning and retention of concepts: IV. The influence of complexity of the stimuli. J. exp. Psychol., 1946c, 36, 252-61.

21. Smoke, K. L. An objective study of concept formation. Psychol. Monogr., 1932, 42, No. 191.

22. Spence, K. W. The nature of response in discriminative learning. Psychol. Rev., 1952, 59, 89-93.

23. Weigl, E. Zur Psycholobie sogenannter Abstraktionsprozesse: I. Untersuchungen uber das 'Ordnen.' Z. Psychol., 1927, 103, 1-45.

24. Wertheimer, M. Productive thinking. New York: Haper, 1945.